

	Match	27.6%;	Score 328.2;	DB 9;	Length 739;
	Local Similarity	99.18;	Pred. No. 3.7e-30;		
	Mismatches	0;	Mismatches	3;	Indels 0; Gaps 0;
QY	801	aaaaggcgcggaagggtctctgcacacgcgcacactggcctgctcccgctccgcgcgcg	860		
Db	2	AAAGGCGCGCGAAGGGGCTCTGCACCGCGCCACTTGGCCTGCCTCCGTCGCCCGCG	61		
QY	861	ccacttggcctgcctcctgcgcgcgcgcacactgcctgcctccgtcccccgcgcgcg	920		
Db	62	CCACTTGCCTGCTCTGCTGCGCGGCCACTTGTGCTGCTCGTCCCGCGCCCGCGCG	121		
QY	921	cgcactgctgtgcgcgtctggagctgcgcgcgcgcctctgcctcccccgcgcacagga	980		
Db	122	CGCCATGCTGTGCGCGGCTGGAGCTGCCGCGCGCGGCGCTTGCCCGCGCGCACAGGA	181		
QY	981	ggggacgcgcgcgcgcgcgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1040		
b	182	CGGGAGCGCCGAGCGCGCTCGCGCCGACGCGGAGCTGCAGTACCTTGGGCGAGATCCAACA	241		
Y	1041	catctccgctgcgcgcgtcgc	1100		
b	242	CATCTTCGCTGGCGGCTGAGAGGACGACCGACGCGGCACCGGCCCTGTCGCTATT	301		
Y	1101	cgc	1133		
b	302	CGGCATGCGCGCGCTACACCTTGAGAGATGA	334		

[illegible]

Query Match 27.6%; Score 328.2; DB 9; Length 819;
Best Local Similarity 99.1%; Pred. No. 3.5e-30;

[illegible]

5
RESULT AL545940
LOCUS
DEFINITION AL545940 LTI_NFL006_P12 Homo sapiens CDNA clone CS0DI023YK20 5 prime, mrna linear EST 16-FEB-2001
ACCESSION AL545940
VERSION
KEYWORDS GI:12878592
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 786)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2003)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
Location/Qualifiers
1. 786
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1023K20"
/clone_lib="LFL.NFL006.Pl2"
/tissue_type="Placenta"
/notes="Vector: pCWSPORT 6; Site1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCWSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fullang@lifetech.com URL :
http://fulllength.invitrogen.com
159 a 233 C 233 g 152 t 9 others
BASE COUNT
ORIGIN

```

Query Match 27.6%; Score 327.8; DB 9; Length 786;
 Best Local Similarity 98.8%; Pred. No. 4e-30;
 Matches 329; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 801 aaaggcgcgcgaagggtctctgcaccgcgcacttggctgcctccgcgcgcg 860

```

FEATURES
source
Location/Qualifiers
1. .985
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:5576962"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
194 a. 323 c 264 g 204 t
BASE COUNT
ORIGIN
Query Match 26.6%; Score 315.8; DB 10; Length 985;
Best Local Similarity 97.9%; Pred. NO. 9.2e-29;
Matches 320; Conservative 0; Mismatches 7; Indels 0; Gaps
QY 807 cgcgcggaagggtctctgcaccgcgcacttggctgctctccgcgcgcgcactt 866

```

```

1. 000
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE006YA08"
/clone_lib="LAI FL002_PL1"
/lab_host="DH10B"
/notes="Organ: placenta; Vector: pCMWSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not
I and cloned into the Not I and Eco RV sites of the
pCMWSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      139 a      208 c      192 g      126 t      3 others
ORIGIN
Query Match      26.1%; Score 310.4; DB 9; Length 668;
Best Local Similarity 99.0%; Pred. No. 4.9e-28;
Matches 311; Conservative 1; Mismatches 2; Indels 0; Gaps
QY      820 tctgtccacgcgcacttgcttcgtccgtccgcgcgcacacttgctgcctcgct 879
Db      1 TCTTGTCCACGCGCCACTTGGCTCGTCCGTCCGCGCGCCACTTGCTGCTCGTCCGCGCGCCACTTGCTGCTCGTCCGCTCGT 60

```

us-09-963-333-7.rst

at Jul 27 13:16:46 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

nucleic - nucleic search, using sw model

on: July 26, 2002, 15:13:16 ; Search time 3006.25 seconds

(without alignments)
5329.193 Million cell updates/sec

le: US-09-963-333-7
fect score: 1187
uence: 1 gatcgccactgcaactca.....gagggagcgcgctgggga 1187

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

rchd: 13736207 seqs, 6748477542 residues

al number of hits satisfying chosen parameters: 27472414

inum DB seq length: 0

inum DB seq length: 200000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vrt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	ID	Description
1	335.8	28.3	1060	BM456638	AGENCOURT
2	329.8	27.8	741	AU118644	
3	328.2	27.6	739	AU143295	
4	328.2	27.6	819	AU143180	
5	327.8	27.6	786	AL545940	AL545940 AL545940
6	315.8	26.6	985	BM466216	AGENCOURT
7	310.4	26.1	668	AL541351	
8	303.8	25.6	938	AL551990	AL551990 AL551990
9	295.8	24.9	738	AU130205	
10	294.2	24.8	526	AU142980	
11	288.6	24.3	526	AU128726	
12	286.8	24.2	529	BE794349	
13	285	24.1	905	BI597722	
14	280.2	23.9	935	AL542409	
15	279.6	23.8	524	AG482405	
16	278.8	23.8	1045	BM478045	

18	277	23.3	832	10	BF685317
19	276.8	23.3	742	9	AL548721
20	275.2	23.3	1020	9	AL547612
21	266.6	22.5	664	10	BI520187
22	266	22.4	963	10	EG337540
23	265.8	22.4	809	9	AU117646
24	265.4	22.4	538	10	BI668571
25	265.2	22.3	888	10	BM413631
26	265	22.3	439	10	BF689564
27	263.8	22.2	672	10	EG574267
28	263	22.2	950	9	AL550812
29	262.8	22.1	688	9	AL543217
30	259.8	21.8	836	10	BI822334
31	258.8	21.8	1533	11	BC018858
32	258	21.7	783	10	BE731577
33	255.4	21.5	932	9	AL528821
34	254.8	21.5	442	10	BE280876
35	254.8	21.5	465	10	EG749145
36	253.8	21.4	454	10	BE389795
37	252.2	21.2	761	10	BE409629
38	251.4	21.2	810	10	BE882773
39	248.2	20.9	1035	10	BM473957
40	246.8	20.8	995	9	AL515911
41	245.8	20.7	932	9	AL551817
42	242.8	20.5	893	10	BM457044
43	240.8	20.3	458	10	EG329382
44	240.8	20.3	854	9	AL556369
45	238.8	20.1	897	9	AL519416

ALIGNMENTS

RESULT 1
BM456638 1060 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6408748 NIH_MGC_85 Homo sapiens cdna clone IMAGE:5496315
DEFINITION 5', mRNA sequence.

ACCESSION BM456638
VERSION BM456638.1 GI:18505678
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1060)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS Unpublished (1999)
TITLE Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgabps@mail.nih.gov

COMMENT Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12124 row: k column: 04
High quality sequence stop: 530.

FEATURES

Location/Qualifiers
1..1060
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5496315"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NOTI;
Site: 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

[illegible]

Db 298 CGCGCTACAGCCTGAGAGATGA 319

Search completed: July 26, 2002, 15:13:23
Job time: 26887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2002, 17:57:13 ; Search time 406.4 Seconds
(without alignments)
5014.708 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccactgcactcca.....gaggaggcgctgggga 1187

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	1163.8	98.0	18596	22	AAF31109	Thymidylate synthase
2	1163.8	98.0	18596	22	AAF31109	Human thymidylate synthase
3	287.2	24.2	3298	24	AAS94945	Human DNA sequence
C 4	206.8	17.4	9519	22	AA137169	Human musculoskele
C 5	206.8	17.4	56632	22	AAK65581	Human immune/haema
C 6	206.4	17.4	9731	22	AA104943	Human reproductive
C 7	205.8	17.3	1539	23	AA584960	DNA encoding novel
C 8	204.6	17.2	11453	22	AAK73272	Human immune/haema
C 9	204.6	17.2	11553	22	AAK73271	Human immune/haema

C 10	204.2	17.2	37449	22	AAK66874	Human immune/haema
C 11	204.2	17.2	56743	22	AAK68202	Human immune/haema
C 12	204.2	17.2	56743	22	AAK81760	Human immune/haema
C 13	203.6	17.2	22601	22	AAK79760	Human immune/haema
C 14	202.6	17.1	22916	22	AAK65305	Human immune/haema
C 15	201.8	17.0	10820	22	ABA16598	Human nervous syst
C 16	201.2	17.0	5591	22	ABA15206	Human nervous syst
C 17	201.2	17.0	5591	22	AAK52300	Genomic sequence #
C 18	200	16.8	9060	24	AAK52300	Human genomic DNA
C 19	200	16.8	9060	24	AAK52300	Human genomic DNA
C 20	199.4	16.8	32169	22	ABA14358	Human nervous syst
C 21	198.6	16.7	32173	22	ABA07767	Human ovarian and
C 22	198.6	16.7	32173	22	AAK03558	Human reproductive
C 23	198.6	16.7	38928	22	AAK68452	Human immune/haema
C 24	197.8	16.7	646	22	AAK44995	cDNA encoding nove
C 25	197.6	16.6	7017	22	AAK37408	Human musculoskele
C 26	197.6	16.6	50000	21	AAK96365	Polymorphic repeat
C 27	197.4	16.6	45546	20	AAK23520	Human kidney amino
C 28	196.8	16.6	97662	22	AAK83908	Genomic sequence o
C 29	196.8	16.6	106746	21	AAK10225	Human PCTA-1 genom
C 30	196.6	16.6	7537	22	AAK36235	Human cardiovascular
C 31	196.6	16.6	7537	22	AAK36235	Human immune/haema
C 32	196.6	16.6	10483	22	AAK80686	Human immune/haema
C 33	196.6	16.6	17357	22	AAK86680	Human immune/haema
C 34	196.6	16.6	17363	22	AAK86681	Human immune/haema
C 35	196.6	16.6	345791	22	AAK43104	Human Oestrogen re
C 36	196.4	16.5	13927	22	AAK97853	Human neuroblastom
C 37	196.4	16.5	32012	22	AAK05864	Human reproductive
C 38	196.4	16.5	32248	22	ABA20412	Human nervous syst
C 39	196.4	16.5	32248	22	AAK37122	Human musculoskele
C 40	196.4	16.5	32248	22	AAK26795	Human genomic DNA
C 41	196.4	16.5	34658	22	AAK69489	Human immune/haema
C 42	196.2	16.5	1376	21	AAK99845	Human secreted pro
C 43	196.2	16.5	2233	22	AAK77256	Human immune/haema
C 44	196	16.5	54548	21	AAK45596	DNA sequence of th
C 45	195.8	16.5	16183	22	AAK74819	Human immune/haema

ALIGNMENTS

RESULT 1
AAF31109
ID AAF31109 standard; cDNA; 18596 BP.
AC AAF31109;
XX AAF31109;
XX AAF31109; (first entry)
DT 27-APR-2001 (first entry)
DE Thymidylate synthase coding sequence.
KW Analyte-binding enzyme; analyte analysis; ss.
OS Homo sapiens.
XX
XX WO200102600-A2.
XX
XX PD 11-JAN-2001.
XX
XX PF 30-JUN-2000; 2000WO-US18057.
XX
XX PR 06-JUL-1999; 99US-0347878.
XX PR 06-DEC-1999; 99US-0457205.
XX
XX (GAT) GEN ATOMICS.
XX
XX PI Yuan C;
XX
XX DR WPI; 2001-071583/08.
XX
XX PT Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226561.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249308.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

PS Disclosure; SEQ ID NO 28083; 307lpp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic


```
QY 342 tcacaccgctcgcccccacaaagtactaggattacagcgtgagccacacggtccacg 401
Db 2189 ttgtcccgcttgacctcccaaaagtgtggattacaggtgacccacatgcccggcc 2248
QY 402 cctggtcggtttttaacaagttagaaaag 430
Db 2249 aagtatcctttttaataacttggataaag 2277

RESULT 13
AAK79760/C
ID AAK79760 standard; DNA; 20601 BP.
XX
AC AAK79760;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34572.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WC200157182-R2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2002, 16:37:26 ; Search time 5000.16 seconds
(without alignments)
4967.801 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgactcca.....gagggagcgcgctgggga 1187

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_cm.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.							

1	1163.8	98.0	1186	9	HUMTS	D00517 Homo sapien
2	1163.8	98.0	18596	6	ARI44965	ARI44965 Sequence
3	1163.8	98.0	18596	6	AX050451	AX050451 Sequence
4	1163.8	98.0	18596	6	AX330682	AX330682 Sequence
5	1163.8	98.0	18596	6	AX330906	AX330906 Sequence
6	1163.8	98.0	18596	6	AX335755	AX335755 Sequence
7	1163.8	98.0	18596	6	HUMTS1	D00596 Homo sapien
8	1002.8	84.5	165742	2	AC020697	AC020697 Homo sapi
9	767.4	64.7	1186	6	AX348477	AX348477 Sequence
10	708.4	59.7	1186	6	AX348476	AX348476 Sequence
11	297.8	25.1	1536	6	ARI03889	ARI03889 Sequence
12	297.8	25.1	1536	6	HSTSVM	X02308 Human mRNA
13	287.2	24.2	3298	6	AX281791	AX281791 Sequence
14	277	23.3	1569	9	BC013919	BC013919 Homo sapi
15	274.2	23.1	1533	9	BC002567	BC002567 Homo sapi
16	228.2	19.2	23454	2	AP005703	AP005703 Homo sapi
17	227.6	19.2	195622	2	AP000807	AP000807 Homo sapi
18	225.2	19.0	80585	9	AL357374	AL357374 Human DNA
19	224.4	18.9	190871	9	AC044797	AC044797 Homo sapi
20	222.6	18.8	202827	9	AC016907	AC016907 Homo sapi
21	221.6	18.7	160536	9	AL355984	AL355984 Human DNA
22	221	18.6	150296	9	HUT91326	U91326 Human Chrom
23	220.6	18.6	97137	9	AC004745	AC004745 Homo sapi
24	220.4	18.6	132948	9	HS349412	AL033520 Human DNA
25	219.6	18.5	146407	9	AC004953	AC004953 Homo sapi
26	219.6	18.5	195773	2	AC073134	AC073134 Homo sapi
27	219.6	18.5	224896	2	AC098591	AC098591 Homo sapi
28	219.4	18.5	104289	2	HS104D14	AL035554 Homo sapi
29	219	18.4	43497	9	HS695020A	AL049856 Human DNA
30	219	18.4	202676	2	HSDA109B7	AL121672 Homo sapi
31	218.8	18.4	156602	9	AC012487	AC012487 Homo sapi
32	218.6	18.4	146954	9	AL590640	AL590640 Human DNA
33	218.2	18.4	180346	9	AC092615	AC092615 Homo sapi
34	218.2	18.4	190268	2	AC108222	AC108222 Homo sapi
35	217.6	18.3	305000	9	HSXDPA	AL590752 Homo sapi
36	216.6	18.2	149546	9	AC087256	AC087256 Homo sapi
37	216.6	18.2	152097	2	AC024225	AC024225 Homo sapi
38	216.6	18.2	196817	9	AC019209	AC019209 Homo sapi
39	216.4	18.2	80914	9	AC007999	AC007999 Homo sapi
40	216.4	18.2	169007	9	AC034193	AC034193 Homo sapi
41	216.4	18.2	178324	9	AC018828	AC018828 Homo sapi
42	216.4	18.2	193214	9	AC069461	AC069461 Homo sapi
43	216.4	18.2	198546	2	AC069337	AC069337 Homo sapi
44	216.4	18.2	207698	9	AC018808	AC018808 Homo sapi
45	216.2	18.2	4118	11	G66726	G66726 csnpcatgic-

ALIGNMENTS

RESULT	1	HUMTS	1186 bp	DNA	linear	PRI 06-NOV-2001
LOCUS	HUMTS	Homo sapiens gene for thymidylate synthase, exon 1, partial cds.				
DEFINITION	D00517					
ACCESSION	D00517					
VERSION	D00517.1	GI:220133				
KEYWORDS		Homo sapiens DNA, clone:lambda HTS-1.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	(bases 1 to 1186)				
AUTHORS		Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T.				
TITLE		Human thymidylate synthase gene: isolation of phage clones which cover a functionally active gene and structural analysis of the region upstream from the translation initiation codon				
JOURNAL		J. Biochem. 106 (4), 575-583 (1989)				
MEDLINE		90110051				
COMMENT		These data kindly submitted in computer readable form by: Keiichi Takeishi				
		University of Shizuoka School of Food and Nutritional Sciences 395 Yada				


```

..021 tacctgggagatccaaacatctccctgctgaggaagagacacgacgagcgcg 1080
|||||
1097 TACTTGGGGAGATCCAAACATCCCTCCGCTGGCGGTGAGAGAGACACGCGACGGCG 1156
|||||
QY 1081 accgacacctgtgatttcgcatgagcgagcgctacagcctgagagtgaccccg 1140
|||||
Db 1157 ACCGGACCCCTGTGCGTATTCGCGATGCGAGGCGCCCTACAGCTGAGAGGTGACGCGCG 1216
|||||
QY 1141 ggcctctgggagcgggtggcggaaggagggcgcgctgggga 1187
|||||
Db 1217 GCGCCCTCGGGAGCGGTGCGGGAAGGAGGAGGCGGCGGTGGGA 1263
|||||

RESULT 5
AX330906 AX330906 18596 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 1415 from Patent WO0194629.
ACCESSION AX330906
VERSION AX330906.1 GI:18103885
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrikan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 1415 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN

Query Match 98.0%; Score 1163.8; DB 6; Length 18596;
Best Local Similarity 99.1%; Pred. No. 3e-191;
Matches 1176; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 gatcgccactgcactccagctcgtggtgagagcgagactctgtcctcaaaaaaaa 60
|||||
Db 78 GATCGCCGCACTGCACCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
|||||
QY 61 aaaaagacgcgaggggtcctcaaaaaaacctcgaaaagccctggcggtcttttttt 120
|||||
Db 138 AAAAAAGCGCGAGGGCTCAACAAAAAACCTCGGAAAGCCCTGGCGGTCTTTT 197
|||||
QY 121 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt 180
|||||
Db 198 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 257
|||||
QY 181 cggatctggctcactgaacctcgtccaggttccaggttccaggttccaggttccag 240
|||||
Db 258 CGGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317
|||||
QY 241 tcccaagtagcacacgcccagcgaatttttttttttttttttttttttttttttt 300
|||||
Db 318 TCCCAAGTAGCACACGCGCGAGCTAATTTTGTACTTTTGTAGTAGAGCGGGTTT 377
|||||
QY 301 catgtgtccaggtgtgtctgataactctgaactcaggtgtatccacccgctcggccc 360
|||||
Db 378 CATGTGTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 437
|||||
QY 361 caaagtactagattacaggggtgagccacgctccagcctcagcctcagcctcagc 420
|||||
Db 438 CAAAGTACTAGGATTACAGGCGTGAAGCCACCGGCTCCAGCGGCTTGTGTGTGT 497
|||||
QY 421 agtagaaaagctgcattataccatttcttcttcttcttcttcttcttcttcttct 480
|||||

```

```

Db 498 AGTAGAAGAGCTGCATTATACCACTTCTTGGTTGC-TTCAGTGAGAACGAAATGG 556
|||||
QY 481 aaatgcaaatAnottattatttagtaaaacagatctcaaacacagttttgtngacaag 540
|||||
Db 557 AAATGCAAAATCCCTTATTAGTTGTAGAAAACAGATCTCAAAACAGACAGTTTGTG 616
|||||
QY 541 accgaggaacacgtgggaactgtctgcttagaagagcgcggtgacacagcg 600
|||||
Db 617 ACAGCAGGAAACGTTGGAACTGTGCTGTAGAGAGGCGCGTGCACACAGCG 676
|||||
QY 601 ttcccaagggcgagtccttccnccgacccgacactgcntccaggttcccgggtntct 660
|||||
Db 677 TTCCCAAGAGGCGAGTCTTCCAGCCACGACCTGATCCAGGTCCCGGTTTCT 736
|||||
QY 661 aagactctcagctgtggccctgggtcttctgtgcaacacccgtggtctcgtcttct 720
|||||
Db 737 AAGACTCTCAGCTGTGGCTGGCTCCGCTTCTGTGCCACACCGCTGCTCTCTGCTTTC 796
|||||
QY 721 cccctggcgacgctctctctagagggggcgcgcgagcccgccgacccgacgagagcg 780
|||||
Db 797 CCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGGAGGAGAGGCG 856
|||||
QY 781 gacgcggagcgcgcggaagagcgcgcggaaggggtctctccaccgcgcacttgg 840
|||||
Db 857 GAGCGGGAGCGCGCGGGAAGAGCGCGGAGGGGTCTCTGCCACCGCGCCACTTGG 916
|||||
QY 841 cctgctcctcctcgcgcgcgcacttggcctcctcctcctcctcctcctcctcctc 900
|||||
Db 917 CTGCTCTCGTCCGCGCGCCACTTGGCTGCTCTGCTCCGCGCGCGCTTCCGCTG 976
|||||
QY 901 cctcctcctcctcgcgcgcgcacttggcctcctcctcctcctcctcctcctcctc 960
|||||
Db 977 CTCTGCTCTCCGCGCGCGCGCATGCTGCTGGCGGCTCGGAGCTCCGCGCGCGGCC 1036
|||||
QY 961 Ctgcctcctcctcgcgcgcgcacttggcctcctcctcctcctcctcctcctcctc 1020
|||||
Db 1037 TTGCCCCCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
|||||
QY 1021 tactgggagcagatccaaacatctcctcctcctcctcctcctcctcctcctcctc 1080
|||||
Db 1097 TACTTGGGCAATCCACATCTCTCCGCTGGGCGTTCAGAGAGGAGGAGGAGGAGGAG 1156
|||||
QY 1081 accgacacctgtcgttattgtgagcgagggcgctacagcctgagagtgacgacg 1140
|||||
Db 1157 ACCGGACCCCTGTGCGTATTCGCGATGCGAGCGCGCTACAGCTGAGAGGTGACGCGCG 1216
|||||
QY 1141 ggcctctgggagcgggtggcggaaggagggcgcgctgggga 1187
|||||
Db 1217 GCGCCCTCGGGAGCGGTGCGGGAAGGAGGAGGCGGCGGTGGGA 1263
|||||

RESULT 6
AX335755 AX335755 18596 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 6264 from Patent WO0194629.
ACCESSION AX335755
VERSION AX335755.1 GI:18126474
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrikan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6264 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers 1..18596
/db_xref="taxon:9606"
Source

```

```
/protein_id="BAA00472.1"
/db_xref="GI:220136"
/translation="MPVAGSELPRRLPPAAQERDABRPBPHGLQVLGOIHILRCG
VKRDRRTGGLSVFGMOARSLRDEPLITTKRVKGVLELLAFIKGSDNAKELS
SGVKIWRANGSROFLDSLGFSSTREGLDLPVYGQWRHGAERYMDESYSQGVQD
LQRIIDITKTPDDRIIMCAWNPRLPLMALPPCHALCOFYVYNSELSCOLYQSGD
MGLGVFFNIASYALITYIAHITGLKPGDFIHTLGDHAHYILNHIETPLKIQLOREPFP
PKLIRLRVEXIDDFKAEDFQIEGYNPHTIKNEMAV"
intron
1206..2894
/number=1
protein_bind
1660..1665
/bound_moiety="Spl"
protein_bind
1890..1895
/bound_moiety="Spl"
protein_bind
1895..1900
/bound_moiety="Spl"
exon
2895..2968
/number=2
intron
2969..5395
/number=2
repeat_unit
3049..3347
/notes="Alu sequence"
repeat_unit
4953..5245
/notes="Alu sequence"
exon
5396..5570
/number=3
intron
5571..11842
/number=3
repeat_unit
5768..9405
/notes="11 sequence"
repeat_unit
10125..10389
/notes="reiterated motif of GGA/TGAT"
exon
11843..11944
/number=4
intron
11945..13448
/number=4
repeat_unit
12135..12397
/notes="Alu sequence"
repeat_unit
12567..12743
/notes="Alu sequence"
repeat_unit
12807..13084
/notes="Alu sequence"
exon
13449..13624
/number=5
intron
13625..14132
/number=5
repeat_unit
13875..14010
/notes="Alu sequence"
exon
14133..14204
/number=6
intron
14205..15612
/number=6
repeat_unit
14531..14832
/notes="Alu sequence"
exon
15613..15750
/number=7
polyA_signal
15934..15939
polyA_signal
16228..16233
repeat_unit
17262..17553
/notes="Alu sequence"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.0%; Score 1163.8; DB 9; Length 18596;
Best Local Similarity 99.1%; Pred. No. 3e-191;
Matches 1176; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 1 gatcgccactgcactccagctgggtgagagcgagactctgtctcaaaaaaaaaa 60
Db 78 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 aaaaagacgcgcagggctcaacaaacacctcggaagacccctggcggtctttttt 120
```

QY 391 qcgtccagcgccctggcggttttttaatacaagtagaagaaagctcattatcacacttgctt 450
|||||
Db 19353 CGCGTCAGCGCCCTGGCGGTTTTTAATCAAGTAGAGAAAGCTGCATATACACATGCGTT 19294
QY 451 cngttgonttoagtgagaaacgaagaaatggaatgcataatattattagtttaggaaa 510
|||||
Db 19293 CGGTTGCGCTTCAGTGAGAACGAGAAATGGAATGCAATCCCTATTAGTTGTAGGAAA 19234
QY 511 cagatctcaaacagagttttttttnacaaagaccagaaagaaagctggagactgtgctgc 570
|||||
Db 19233 CAGATCTCAACAGCAGTGTTCGACAGAGACCGCAGAGAAAGCTGGGAACGTGCTGCT 19174
QY 571 gcttagagaagcgcggtcgaccagaggttcccaaaagggcgagctcttccngccac 630
Db 19173 GGCITAGAGAGCGCGGTGACAGAGCGTTCCTCAAGGGCGCAGCTCTCCCGCGCAC 19114
QY 631 cgcactgonttoagtttcccggttntcttaagactctcagctgtgctgggctcgt 690
|||||
Db 19113 CGCAGCTGCATCCAGGTTCCCGGGTCTCCTAAGACTCTCAGCTGTGGCGCTGGGTCGCT 19054
QY 691 tctgtgcacacccggtgctcgtggtttccctggtgctgacgctctctagagcggggc 750
Db 19053 TCTGTGCCACACCGGTGCTGCTGCTTCCCTCTGGCGCACGCTCTCTAGAGCGGGGC 18994
QY 751 qcgcgcaccccgcgagcaggaagcgagcgcggggagcgccgcgggaaagcgcg 810
Db 18993 CGCGCGAGCCCGCGAGCAGGAGCGGAGCGCGGGAGCGCCGCGGAAAGCGCG 18934
QY 811 qggaaggggtctcyccacgcgcacttggcctcctcccgcgcgccacttgcc 870
Db 18933 CGGAAGGGGTCTTGCCA-----CGCGCGCACTTGCC 18902
QY 871 tqcctcgtcccgcgcgcaacttgcctgctcgttcccccgcgcgcgcgcgcgcgcgc 930
Db 18901 TGCCTCGTCCCGCGCGCCACTTGCCTGCTCCTGCTCCCGCGCGCGCGCCATGCT 18842
QY 931 qggcgcgctcgagctcgcgcgcgccctcgcccccgccgcacagagcgagcgcc 990
Db 18841 GTGGCGGCTCGAGTGGCGCGCGCCCTTGGCCCCCGCGCACAGAGCGGACGCC 18782
QY 991 gagcgcgctcgcgcgagctcgagctacgtgggagatccatccacacacatccgc 1050
Db 18781 GAGCGCGTCCGCGCGAGCGAGCTGAGTACCTGGGGCAGATCCACACATCTCGC 18722
QY 1051 tqcgcgctcaggaagcagcagcgcgagcgcccgcccgcccgcccgcccgcccgcc 1110
Db 18721 TCGCGCGTTCAGGAAGCAGCAGCGCAGCGGCGACCGCCCTGTCGGTATTCGGCATGCG 18662
QY 1111 qcgcgtacagctcgagagtgagcgcgccctcgcgcgccctcgcgcgaggtgagcgag 1170
Db 18661 GCGCGCTACAGCTGAGAGTGAGCGCGCGCGCCCTTGGCGGACGCGGTGGCGGAGGAG 18602
QY 1171 ggaagcgcgctgggga 1187
Db 18601 GGAGCGCGGCTGGGA 18585

RESULT 9
AX348477/c 1186 bp DNA linear PAT 06-FEB-2002
LOCUS AX348477
DEFINITION Sequence 172 from Patent WO0202806.
ACCESSION AX348477
VERSION AX348477.1 GI:18614513

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
1 (sites)

REFERENCE
Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL Patent: WO 0202806-A 172 10-JAN-2002;
Epigenomics AG (DE)

FEATURES

source Location/Qualifiers
1. .1186
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 229 a 98 c 386 g 473 t
ORIGIN
Query Match 64.7%; Score 767.4; DB 6; Length 1186;
Best Local Similarity 78.4%; Pred. No. 6.2e-123;
Matches 926; Conservative 0; Mismatches 254; Indels 1; Gaps 1;
QY 2 atcgcgccactgcactccagctgggtgagagcgagagctgtctctcaaaaaa 61
Db 1185 ATCGCGCCACTACACTCCCACTAAATAAAAAACGAAACCTCTATCTCAAAAAA 1126
QY 62 aaaaacgcgcaggggtcnaaaaaaaacccctggaaagccctggcggtcttttttt 121
Db 1125 AAAAAACGCGCAAACTCAACAAAAAACCTCGAAAAAACCTTAACGATCTTTTTTT 1066
QY 122 ttt 181
Db 1065 TT 1006
QY 182 ggatcttggtcactgcgaactctgctcccgaggttccaagcaattcttctgctcagct 241
Db 1005 GAATCTTAACCTACACTCACTCTACCTCCCAATTCAAAACAATTTCTTCTACCTCA 946
QY 242 ccaagtagtgcacacgc 301
Db 945 CCAATAAACCACGACGCCCACTAAATTTTATATTTTATATAAAACGAAATTTTACC 886
QY 302 agttgtccaggtggtctngaaactcctgaactgaactgaactgaactgaactgaact 361
Db 885 ATATTATCCAACTAAATCTTAAACTCTTAACCTCAATTAATTAATTAATTAATTA 826
QY 362 aaagtactagattacagcggtgagcgcccgctccagcgccctggcggttttttaata 421
Db 825 AAATACATAAATTAACAAAGTAACACCGCGCTCCACGCGCTCAACGATTTTATCAA 766
QY 422 gtgaaaaagtgcatattacacttgcttongtngtnttcagtgagaacgaagaatgga 481
Db 765 ATAAAAAATACATATATACACTTACTTCGATAC-TTCAATAAAAAAGAAAAATAA 707
QY 482 aatgcaaatantctattagtttaggaaacagattcctcaaacagcagttttgtngaca 541
Db 706 AATACAAATCCCTTATTAATTAATAAAAAACAAATCTCAACAAACAAATTTATTA 647
QY 542 cgcagggaaaacgltgggaactgtgctggttagagaagcgcggtcgaccagacggt 601
Db 646 CCGCAAAAAACGTAACAACTATACCTAACTTAATAAAAAACGCGATCGACCAACGAT 587
QY 602 tcccaagggcgagctctcccgccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 661
Db 586 TCCCAAAAAACGCAATCTTCCCAACCGCACCTACATCAAAATTCGCGAATTCCTA 527
QY 662 agactctcagctgtggccctggcctcgtttgtgcaaccccggtggtcctgctgttcc 721
Db 526 AAACCTCAACTATTAACCTTAACTCCGTTCTATACACACCGCTAACCTCTACGTTCC 467
QY 722 ccttcgacagctctctagagcgggcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 781
Db 466 CCTAACCGACGCTCTCTAAACGAAACCGCGCGGACCGCGCGGACCAAAAAACA 407
QY 782 agcgcgagcgccgcgggaaagcgcgcggaaggggtcctgccaccgcgcacttgcc 841
Db 406 AACGGAAACGACCGGAAAAAAGCGCGGAAAAAATCTTACCACCGCGCACTTAAC 347
QY 842 ctgctcctcgcccgcgccacttggcctcctcctcctcctcctcctcctcctcctcct 901
Db 346 CTACCTCGTCCCGCGCGCACTTAACCTACCTCGTCCGCGCGCGCGCGCGCGCTAC 287

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2002, 14:19:11 ; Search time 77.68 Seconds
(without alignments)
4857.017 Million cell updates/sec

Title: US-09-963-333-6
Perfect score: 1536
Sequence: 1 ggggggggggggaccacttg.....ataataaagaagtgtctgc 1536

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
- 4: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
- 5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq.*
- 6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description
1	1531.2	99.7	1536	3	US-09-089-195-1
2	613.2	39.9	18596	4	US-09-318-448-11
3	387.4	25.2	1014	4	US-09-230-637-1
C 4	387.4	25.2	35100	2	US-08-770-379-17
C 5	387.4	25.2	35100	4	US-08-757-669A-17
C 6	387.4	25.2	35100	4	US-09-230-371A-17
C 7	121.2	7.9	3479	3	US-08-714-918-7
C 8	121.2	7.9	3479	4	US-09-265-315-7
C 9	121.2	7.9	3479	4	US-09-265-315-7
C 10	121.2	7.9	3479	4	US-09-266-417-7
C 11	48.6	3.2	503	4	US-08-483-533-7
C 12	48.6	3.2	503	4	US-09-283-471A-7
C 13	48.6	3.2	1292	4	US-08-483-533-37
C 14	48.6	3.2	1292	4	US-09-283-471A-37
C 15	47.6	3.1	702	1	US-08-458-568A-3
C 16	47.6	3.1	12001	1	US-08-458-568A-11
C 17	46.4	3.0	152331	3	US-09-128-155-16
C 18	46.4	3.0	4403765	4	US-09-103-840A-2
C 19	46	3.0	1335	5	PCT-US91-06532-1
C 20	46	3.0	4488	4	US-08-406-030A-3
C 21	45.8	3.0	1550	2	US-08-609-443B-17
C 22	45	2.9	320	4	US-09-165-264-7
C 23	45	2.9	1300	4	US-08-483-533-39
C 24	45	2.9	1300	4	US-09-283-471A-39
C 25	45	2.9	1327	4	US-08-483-533-36
C 26	45	2.9	1327	4	US-09-283-471A-36
C 27	44.6	2.9	44377	2	US-08-804-227C-7

28 44.6 2.9 44377 2 US-08-804-198-1 Sequence 1, Appli
29 44 2.9 212 4 US-08-483-533-18 Sequence 18, Appli
30 44 2.9 212 4 US-09-283-471A-18 Sequence 18, Appli
C 31 44 2.9 319 4 US-09-165-264-8 Sequence 8, Appli
32 44 2.9 595 4 US-08-483-533-4 Sequence 4, Appli
33 44 2.9 595 4 US-09-283-471A-4 Sequence 4, Appli
34 43.6 2.8 320 4 US-09-165-264-13 Sequence 13, Appli
35 43.4 2.8 1280 4 US-08-483-533-38 Sequence 38, Appli
36 43.4 2.8 1280 4 US-09-283-471A-38 Sequence 38, Appli
C 37 42.8 2.8 320 4 US-09-165-264-14 Sequence 14, Appli
38 42.4 2.8 212 4 US-08-483-533-13 Sequence 13, Appli
39 42.4 2.8 212 4 US-09-283-471A-13 Sequence 13, Appli
C 40 42.2 2.7 320 4 US-09-165-264-11 Sequence 11, Appli
41 42.2 2.7 9757 1 US-08-093-453B-1 Sequence 1, Appli
42 42.2 2.7 9759 1 US-08-459-041A-1 Sequence 1, Appli
43 42.2 2.7 9759 3 US-08-999-733-1 Sequence 5, Appli
44 42 2.7 623 4 US-09-043-303-5 Sequence 1, Appli
C 45 42 2.7 8438 1 US-07-945-283-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-089-195-1
; Sequence 1, Application US/09089195
; Patent No. 6087489
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION
; TITLE OF INVENTION: OF HUMAN THYMIDYLATE SYNTHASE EXPRESSION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,195
; FILING DATE: herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 810-1515
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; PUBLICATION INFORMATION:
; AUTHORS: Takeishi, K.
; AUTHORS: Kaneda, S.
; AUTHORS: Ayusawa, D.
; AUTHORS: Shimizu, K.
; AUTHORS: Gotoh, O.
; AUTHORS: Seno, T.
; TITLE: Nucleotide sequence of a functional cDNA
; TITLE: for human thymidylate synthase
; JOURNAL: Nucleic Acids Res.
; VOLUME: 13
; ISSUE: 6


```

; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-757-669A-17

```

Query Match 25.2%; Score 387.4; DB 4; Length 35100;
Best Local Similarity 63.5%; Pred. No. 1.1e-92;
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

QY	115	gcggctcgagtgccgcyccgycgcccttgcgcccccgcacagagcgggacgcgcgaq	174
Db	21023	GCAGTCTGGCCTTTGGCCCGGAGGTTCACCCCTTGTCGATTCAGCATGASGTAACCTGG	20964
QY	175	ccgcgtccgcgcacggggagctgcagtacctgggcagatccaacatactccgtcgc	234
Db	20963	GCTGAGACGCCCCNACAGAGACTTCAGTAUCTCAGGCAGTTGAGGGAATTTGTGCCGT	20904
QY	235	ggcgtcaaggaaagcacacgcgcgcggaccgcgcacttcgcgtattcggcatgcacggcg	294
Db	20903	GGCACGCATCGACTCGACCGCACCGGCATAGGCACCTCTCTATTGTGGCATCGAGGC	20844
QY	295	cgtcaacgctgagagatatgaattccctcgtgcacaaccaaacacgttgtctctggaagggt	354
Db	20843	AGGTATAGCTCTGGCGGACACACTTCCCTTTACTAACCAAGCGGGTGTTTGGCAGGC	20784
QY	355	gttttgagagttgctgtgtttatcaaggatccaacaaatgctaaaagactgtcttc	414
Db	20783	GTAGTGCAGAGCTGCTTGTGTTTTGAAAGGGGAGTACTGACTCCAGGAGCTTTCAGA	20724
QY	415	aaggagtgaaactctgatgcacaatggatccccagactttttggacagcctgggttic	474
Db	20723	ACAGGAGTCAAGATATGGGACAAAATGGCTCCAGGSAGTTCCTGGCGGCGCGCCGTG	20664
QY	475	tccaccagagaaggggacttgygcccagtttatgtgcttoaatgtaggacattttggg	534
Db	20663	GGCGACAAGGGAGGGGATTTGGGACCTGTTTACGGTTTCCAGTGGAGACACTTTGGG	20604
QY	535	gcagaatacacagatatgaatacagattaticcagacagaggagtgcacaactgcacaga	594
Db	20603	GGGGCGTAGTGGACGGGATGTGACTATATACGGCCAGGGGTTTGACCAATTCGTCTAC	20544
QY	595	gtgattgacacatcaaaacccacctgcacacagaagaatcatactgctgcttggaaat	654
Db	20543	ATTGTGGATTTAATAAAAAATAATCCGCACGATAGAGAATCATTTATGTGCGGTGAAC	20484
QY	655	caaagagatctctctgatggcgtgctccatgcocatgcocctcgtccagttctatgtg	714
Db	20483	CGGCGEGACTTGTCTGTATGGCGCTCCGCCCTGCTACTTGTATGTCAAATTTATGTA	20424
QY	715	gtgaacagtgtgctctccagctgtaccagagatcgggacacatgggacctcggtgtg	774
Db	20423	GCTGACGGTGAAGCTTCCGTCTCAGCTGTATCAGAGGTCGGGACATGGSGTTTGGAGTT	20364
QY	775	cotttcaacatcgcgcagctaogcctgtcactacatgattgcgcacatcacogggocgtg	834
Db	20363	CCTTTTAACATTGCCAGCTATTCCCTCTTAATATGTTGCTCATGTTACTGGTCTT	20304
QY	835	aagcgaagtgcactittatacaactttgggaagatgcaaatatttaacctgaatcacatcgag	894
Db	20303	AGACOCGGGGAGTTATTTCACGTTGGGAGATGCCACATCTACAAAACGATATAGAG	20244
QY	895	ccactgaaaaatcagcttcagcagaacccagacactttccaaagctcaggtattctcga	954
Db	20243	CCACTACGGCTGCAGCTGACGCGCATCCAGTCCCCTTCCGGGCTGGAGATCTCCGG	20184
QY	955	aaagttagaaaaattgatgacttcaaaagctgaagactttcagattgaagggttacaatccg	1014

Db	20183	TCGTGTTCTCCAGGAAGAGATTTACACCTGATGATTTTAGACTGGTGGACTACTGCCCG	20124
Qy	1015	catccaactattaaaatggaaatgcgtgtttag	1047
Db	20123	CATCTACCATTCGTATGGAATGGCAGTATAG	20091

RESULTS 6

```

US-09-230-371A-17/c
; Sequence 17, Application US/09230371A
; Patent No. 6348586
;
; GENERAL INFORMATION:
;
; APPLICANT: Chang, Yuan
; APPLICANT: Borenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
;
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
;
; FILE REFERENCE: 45185-G-PCT-US
;
; CURRENT APPLICATION NUMBER: US/09/230.371A
;
; CURRENT FILING DATE: 1999-11-17
;
; PRIOR APPLICATION NUMBER: PCT/US97/13346
;
; PRIOR FILING DATE: 1997-07-22
;
; NUMBER OF SEQ ID NOS: 30
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 17
;
; LENGTH: 35100
;
; TYPE: DNA
;
; ORGANISM: kaposi's sarcoma-associated herpesvirus
;
; US-09-230-371A-17

```

[illegible]

RESULT 13
US-08-483-533-37
; Sequence 37, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Rolzman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; TITLE OF INVENTION: Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James F.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..702
US-08-458-568A-3

Query Match 3.1%; Score 47.6; DB 1; Length 702;
Best Local Similarity 53.8%; Pred. No. 0.0027;
Matches 98; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 11 ggaccacttgctgcgtccgcgcgcgcacttggcctgcctccgctccgcgcgcgc 70
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 GGCCCCCACCAGCGCGGTGGGGCCCGGGGGGGGGGGTGAACCCCTCCACCCCCCTCGC 305
QY 71 cacttgcctgcctgcctcccccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 GCGCCCTTCGCGCTTCGCGCGCGCGCTCGCCCTCCGCTCGCGCTCACCGGAGACCTGG 245
QY 131 cgcgcgcgccttgc 190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 CGCGCTTCGCGCTTCGACGCGCGGGGAGAGGGGGCGCGCGAGCCCCCGACCCCCG 185
QY 191 gg 192
Db 184 CG 183

Search completed: July 26, 2002, 17:46:36
Job time: 12445 sec

XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPT; 2001-639362/73.
 DR P-PSDB; ABG20773.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 1; SEQ ID No 20764; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1539 BP; 396 A; 379 C; 347 G; 417 T; 0 other;

Query Match 83.4%; Score 1281.4; DB 23; Length 1539;
 Best Local Similarity 97.4%; Pred. No. 0;
 Matches 1410; Conservative 0; Mismatches 21; Indels 16; Gaps 10;

QY 106 atgctgtggcggctcgagtcgcccgcggcccttgccccgcgcacagagcgg 165
 DB 1539 ATGCTGTGGCGGCTCGAGCTCGCGCGCGGCCCTTGCCCCCGCGACAGGCGG 1480
 QY 166 gaocgcgagccgctccgcgcacggggagctgcagtaacctggggcagatccaaacatc 225
 DB 1479 GACGCGAGCGCGCTCGCGCGCGGAGCTGCAGTACCTGGGCGACATCCAAACATC 1420
 QY 226 ctccgctggcgctcgaaggagacgcgcacgggcacggcgcacctgcgtattcggc 285
 DB 1419 CTCGCGTGGCGGCTCAGGAAGAGAGCGCGACGCGGCGACCGCCCTGCGGTATTCGGC 1360
 QY 286 atgcagcgctcacgctcagagatgaattccctctcgtcacaccaaactgtgttc 345
 DB 1359 ATGCAGCGCGCTCAGAGCTGAGAGATGAATCCCTCTGTCGACAAACAAACGCTGTTTC 1300
 QY 346 tggaggggtttttgagagtgctgtgtttatcaaggagatccacaaatgctaaagag 405
 DB 1299 TGGAGGGTCTTTTGGAGGAGTGTGTGTGTATCAAGGATCCACAAATGCTAAAGAG 1240
 QY 406 ctgcttccaaaggagtgaaatctggagtgccaatggatggatcccgagacttttggacagc 465

DB 1239 CTGCTTCCAGGGAGTGAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGC 1180
 QY 466 ctggattctccaccagagaagagggacttggcccaagttatggcttccagtgagag 525
 DB 1179 CTGGGATTTCCACCAGAGAAAGGGGACTTGGGCCAGTTATATGGCTTCCAGTGAGG 1120
 QY 526 catttggggcagaatacacagagatatgaatcagattattcaggacagggagttgaccaa 585
 DB 1119 CATTTTGGGCGAGATACAGAGATATGAATCAGATTATTCAGGACAGGAGTTGACCAA 1060
 QY 586 ctgcaagagtgattgacccatcaaaacccctgacgcagacagaagaatcatcatgtgc 645
 DB 1059 CTGCAAGAGCTGNTTGCACACATCAAAACCAACCTGACGACAGAGAATCATCATTTGC 1000
 QY 646 -gcttgaatccaagagatc-ttccctctgatggcgctgcctccatgcctcgcctctgc 703
 DB 999 GGGCTGAATCCAAAGAGATCTTCTCTGATGGCGTCCCTCATGCGATGCCCTTGCC 940
 QY 704 agttctatgtgtgaacagt-gagctgctcctgcacgtgta-ccagagatcggagacat 761
 DB 939 AGTTCATGTGTGAACAGTGGAGCTGTCTTCCAAACCTGACCCGGGGATCGGAGACAT 880
 QY 762 gggcctcggtgcttccaacatcgccagctacgcccctgctccagtcacatgattgcga 821
 DB 879 GGGCTGGGTGTGCTTTCACATCGCCAGCTAGGCCCTGCTCAGTCAATGATCGCGCA 820
 QY 822 catcacgggc-ctgaagccagtgtaattatacacactttgggagatgcacatatattacc 880
 DB 819 CATCACGGCTCTAAAGCTAGGTGTCTATCCACACTTGGGGAGATCCACATATTTTAC 760
 QY 881 tgaatcacatcagccactgaaattcagcttcagcgag--aacccagaccttcccaa 938
 DB 759 TGAATCACATCGAGCCACTGAAATTCAGCTCAGCGAGAAACCCAGCTTTCCCAA 700
 QY 939 gctcagattctcgaagttgaaaaattgatgacttcaagctgaagacttcaagat 998
 DB 699 GCTCAGGATCTCTGAAAGTTGAGAAAATGATGACTTCAAGCTGAAGACTTTCAGAT 640
 QY 999 tgaagggtacaatccgcacatcccaactattaaatggaaatggctgtttaggggtcttcaa 1058
 DB 639 TGAAGGTACAATCCGCATCCCACTAATAATGAAATGGTGTCTAGGTGCTTTCAA 580
 QY 1059 aggsagctcgaagatattgtcagctctttaggggttgggtggatccgaggtgaaagtcc 1118
 DB 579 AGGAGCTCGAAGGATATTGTGCTAGTCTTTAGGGGTGGGGTGGATCCCGAGTAAAGTTC 520
 QY 1119 ttttgcctctaaaaagaagaagaaactaggtcaaaaatctgctcgcgcacctcagttatt 1178
 DB 519 TTTTGTCTTAAAGAAAGAAAGAAAGAACTAGGTCAAAAATCTGTCCTGACCTATCAGTTATT 460
 QY 1179 aatttttaaggat-gtggcactggcaatgtactgtgccaagtcttcttccataataaaa 1237
 DB 459 AATTTTAAAGATGGTGGCAATGTAACTGTGCCAGTTCTTTCCATATAAATA 400
 QY 1238 ggctttagttaaactcaactcaggggtatctgacaatgctgaggttatgaacaaactgagga 1297
 DB 399 GGCCTTGAGTTACTACTGAGGGTATCTGACAATGCTGAGTCAATGAACAAAGTGAAGA 340
 QY 1298 gaatgaatgtatgtgctcttag-caaaaactgtatgtgcatttcaatccacgcgtactt 1356
 DB 339 GAATGAATGTATGTGCTCTTAGCCAAAACATGATGATGTGCTATTTCAATCCACGTACTT 280
 QY 1357 ataaagaa-ggttgggtgaattccaacagctatttttgggaatttttgaataattttaag 1415
 DB 279 ATAAAGAGGGTGTGGTGAATTTCAAGCTATTTTGGGAATATTTTGAATAATTTAAG 220
 QY 1416 aatttcaaggtattccctcaaatctbgagggagctbgagtaaacacacatcgatcatgt 1475
 DB 219 AATTTCACAAGCTAATCCCTCAAACTGAGGGAGCTGAGTAACCCCATCGATCATGTAT 160
 QY 1476 agagtgtggttatgaacttaa-----agttgttttatgttgcctataataaagaat 1529
 DB 159 AGAGTGGTGTGAACCTTAAAGTTATAGTTTGTATATGTGCTATATAAAGAAGT 100

ID	AAV73802 standard; DNA; 35100 BP.
XX	
AC	AAV73802;
XX	
DT	25-FEB-1999 (first entry)
XX	
DE	KSHV LJR DNA (nucleotides 1-35,100).
XX	
KW	Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
KW	dihydrofolate reductase; LJR; long unique region; vaccine; prophylaxis;
KW	diagnosis; treatment; HHV8; complement binding protein; v-CBP; SSBP;
KW	sDNA binding protein; transport protein; glycoprotein B; pol; vIL-6;
KW	DNA polymerase; viral interleukin-6; BHV4-IE1 I; thymidylate synthase;
KW	WMIP-II; BHV4-IE1 II; WMIP-I; capsid protein I; tegument protein I; ds.
XX	
OS	Kaposi's sarcoma-associated herpesvirus.
XX	
PN	US5849564-A.
XX	
PD	15-DEC-1998.
XX	
PF	29-NOV-1996; 96US-0770379.
XX	
PR	29-NOV-1996; 96US-0770379.
XX	
PA	(UTCO) UNIV COLUMBIA NEW YORK.
XX	
PI	Bohenzky RA, Chang Y, Edelman IS, Moore PS, Russo JJ;
XX	
DR	WPT; 1999-069741/06.
XX	
PT	Kaposi's sarcoma-associated herpes virus nucleic acid - encodes
PT	dihydrofolate reductase and is useful for treatment, prophylaxis
PT	or diagnosis of Kaposi's sarcoma
XX	
PS	Disclosure; Column 67-96; 109pp; English.
XX	
CC	This sequence is a fragment of the Kaposi's sarcoma-associated
CC	herpesvirus (KSHV) LJR (long unique region). This fragment contains
CC	coding regions for KI, ORF4 which encodes the complement binding protein
CC	v-CBP, ORF6 which encodes a ssDNA binding protein (SSBP), ORF7 which
CC	encodes a transport protein, ORF8 which encodes glycoprotein B, ORF9
CC	which encodes DNA polymerase (pol), ORF10, ORF11, K2 which encodes viral
CC	interleukin-6 (vIL-6), ORF02 which encodes dihydrofolate reductase
CC	(DHFR), K3 which encodes BHV4-IE1 I, K5 which encodes BHV4-IE1 II, K6
CC	synthase, K4 which encodes WMIP-II, K7, ORF16 which encodes Bcl-2, ORF17 which encodes
CC	capsid protein I, ORF18 and ORF19 which encodes Tegument protein I.
CC	KSHV is a new human Herpesvirus (HHV8) believed to cause Kaposi's sarcoma
CC	(KS) which is the most common form of neoplasm occurring in persons with
CC	acquired immune deficiency syndrome (AIDS). The DHFR protein is useful
CC	for vaccination, prophylaxis, diagnosis and treatment of a subject with
CC	Kaposi's sarcoma and for detecting expression of a DNA virus associated
CC	with Kaposi's sarcoma in a cell.
XX	
SQ	Sequence 35100 BP; 8703 A; 9395 C; 8921 G; 8081 T; 0 other;
XX	
Query Match 25.2%; Score 387.4; DB 20; Length 35100;	
Best Local Similarity 63.5%; Pred. No. 2.7e-90;	
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;	
QY	115 gccggctcgagcgtccgcgcccgttgcgccccgcgcacagagcgggacgcgcag 174
Db	82 ggcgtcttgcttgcccccagggtgaacctgttcgatcagcatgagttaactggg 141
QY	175 ccgcgtccgcgcacgggagctcagttaccctgggcagatccaacaattccctgc 234
Db	142 gctgagacgccccacagaaactcagtatctcagcagcttgaggaaatttgtccgt 201
QY	235 ggctcagggaagcacacgcgcacggcaccgcgcacccctgcgttatccgcacggcg 294
Db	202 ggcagcagatcagctcagccgcacggcagcaggtacagcaccctctctatttggatcagggcc 261
QY	295 cgcctacagcctgagatgaattccctctctgtgcacaccaaactgtgtctctggaagggt 354
Db	262 aggtatgctcgcggaccaacttcctactaaccaacagaagggtgttttgcggaggc 321
QY	335 gttttgagagtgctgtgtttatacaaggatccaacaaactctaagagactctcttc 414
Db	322 gtatgcaagctgtctgtgtttctgtaggggagttactgaactccagggagcttccaaga 381
QY	415 aaggagtgaaaatctggatcccaatgatcccgagactttttgacagcctgggattc 474
Db	382 acaggagtcaagatacgggacaaaaatggccccaggagttcctgcgcggcgccgtg 441
QY	475 tccaccagagaagaagggaactggccagcttatgttcctcagtgaggcattttggg 534
Db	442 gcgcacagaaggagggatttgggaactgtttaacgtttccagtgaggcactttggg 501
QY	535 gcagaatacagagatatggatcagattattongacagcggggagttgacaaactgcгааага 594
Db	502 ggccggactcgtgacgcggagctgactatacagccaggggtttgacaaatttcgtac 561
QY	595 gtgattgacaccataaaacaaacccctgcacagagaagaatacatatgtgccttgaat 654
Db	562 attgtggatttaaaaaaatatccgcacgatagaagaatcattatgtgtgctgggaac 621
QY	655 caaagatctctctgtatggctgcctccatgcctatgcctctgcagttctatgtg 714
Db	622 ccggcggactgtcgttgatggcgttcgcgcctgtcaactgttatgtcaattttatgta 681
QY	715 gtgaacgtgagctgtcctgcagctgttacacagagatcggagacatggcctcgtgtg 774
Db	682 gctgacgtgagcttctcgtcagctgtatcagaggtcggagacatgggtttggagtt 741
QY	775 cctttcaacatcgcgcagctacgccctgtccagttacatgattgcgcacatcacggcctg 834
Db	742 ccttttaacatcgcgcagctattccctttaacttatagtgtctcattgtactgtgctt 801
QY	835 aagccagctgactttatacactttggagatgcacatatattacctgaatcaatccag 894
Db	802 agaccocggagttattcaacgcttggagatgccacatctacaaacgpatatagag 861
QY	895 ccaatgaaaaattcagttcagcgagaacccaacgaactttcccaagctcaggaattctga 954
Db	862 ccaactacggctgagctgacgcgcactccacgtccctttccgcgcggagatactccg 921
QY	955 aaagttgaaaaattgatgacttcaagctgaagactttcagattgaaggggtacaatccg 1014
Db	922 tctgtttctccalygaagagtttacactgatattttatagactgtggactactgcocg 981
QY	1015 catccaactattaaaatggaaatggctgtttag 1047
Db	982 catcctacactcgtatggaatgcagttatag 1014
RESULT 10	
AAV73802/C	

/01200CTH /01200CTH

/clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 298 a 253 c 217 g 318 t 3 others
 ORIGIN

Query Match 58.2%; Score 1048; DB 9; Length 1089;
 Best Local Similarity 99.0%; Pred. No. 2.1e-225;
 Matches 1073; Conservative 2; Mismatches 7; Indels 2; Gaps 2;
 QY 432 ggaatgcaatggatccgagacttttgacagcctgggattctccaccagagaagaag 491
 Db 1083 ggaatgcaatggatccgagacttttgacagcctgggattctccaccagagaaga 1026
 QY 492 ggaactgggcccagttatggcttcagtgaggcattttggggcagaatacagagatat 551
 Db 1025 ggacttggccccagttatggcttccagtgaggcattttggggcagaatacagagatat 966
 QY 552 ggaatcagattatccagagcggaggttgaccactcgaagagtgattgacacacataa 611
 Db 965 ggaatcagattatccagagcggaggttgaccactcgaagagtgattgacacacataa 906
 QY 612 aaccaaccctgacagagaagaatacatatgctggcttgaatcccaagagattctcct 671
 Db 905 aaccaaccctgacagagaagaatacatatgctggcttgaatcccaagagattctcct 846
 QY 672 gatggcgtgctccatgcccagtcctcctgcccagttctatgtgtgaacagtgagctgc 731
 Db 845 gatggcgtgctccatgcccagtcctcctgcccagttctatgtgtgaacagtgagctgc 786
 QY 732 ctgcagctgtacagagatcgaggagacatgggootgggtgtgcctttcaacatcgccag 791
 Db 785 ctgcagctgtgtacagagatcgaggagacatgggootgggtgtgcctttcaacatcgccag 726
 QY 792 ctacgcctgtcagctacatgatgctgcacatacaccggcctgaagcaggtgactttat 851
 Db 725 ctacgcctgtcagctacatgatgctgcacatacaccggcctgaagcaggtgactttat 666
 QY 852 acacacttgggagatgcacataattactgaatacactcagccactgaaatcagct 911
 Db 665 acacacttgggagatgcacataattactgaatacactcagccactgaaatcagct 606
 QY 912 tcagcagaacccagactttcccaagctcagatttcttcgaaaaagttgagaataatga 971
 Db 605 tcagcagaacccagactttcccaagctcagatttcttcgaaaaagttgagaataatga 546
 QY 972 tgaactcaaacctgaacacttcaagttgaagttacaatcccatcccatcattataat 1031
 Db 545 tgaactcaaacctgaacacttcaagttgaagttacaatcccatcccatcattataat 486
 QY 1032 ggaatggctgttaggtgtcttcaagaggagctcgaaggatattgtcagctcttagggg 1091
 Db 485 ggaatggctgttaggtgtcttcaagaggagctcgaaggatattgtcagctcttagggg 426
 QY 1092 ttgggtggttcgaggttaaaagtcttttctctaaagaagaagaaactaggtcaa 1151
 Db 425 ttgggtggttcgaggttaaaagtcttttctctaaagaagaagaaactaggtcaa 366
 QY 1152 aaatctgtccgtgacctatcagttattatatttaaggatgttgcactggcaaatgtaa 1211
 Db 365 aaatctgtccgtgacctatcagttattatatttaaggatgttgcactggcaaatgtaa 306
 QY 1212 ctgtgccagttcttccataataaaggcttttgagtttaactcactgaggggtatctgacaa 1271

Db 305 CAGTCCAGTCTCTTCATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAA 246
 QY 1272 tgctgagttatgaacaaagtggagagaatgaaatgtatgtctcttagcaaaaacatgt 1331
 Db 245 TGCTGAGGTTATGAACAAAGTGGAGGAATGAATGATGTGCTCTTAGCAAAAACATGT 186
 QY 1332 atgtgcatttcaatcccaogtacttatataaagaagttgggtgaatttcaacaagctatttt 1391
 Db 185 ATGTGCAATTCATCCACGCTACTTATAAAGAGGTTGGTGAATTTTCAACAAGCTATTTT 126
 QY 1392 ggaatattttagaataattttaagaatttcaacagctattccctcaaatctgagggagct 1451
 Db 125 GGAATATTTTAGAATATTTAAGAATTTCAACAGCTATTCCTCAAACTCAGGGAGCT 66
 QY 1452 ggaataaccatcatcatgtagatgtagagtggtttagaactttaagttgttttatatg 1511
 Db 65 GAGTACACCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6
 QY 1512 ttgc 1515
 Db 5 TTGC 2
 RESULT 3
 AL572496/c
 LOCUS AL572496 LTI_NFL006_PL2 1077 bp mRNA linear EST 16-FEB-2001
 DEFINITION Prime, mRNA sequence.
 ACCESSION AL572496
 VERSION AL572496.1 GI:12930822
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1077)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..1077
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1008YH21"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 294 a 251 c 214 g 309 t 9 others
 ORIGIN

Query Match 67.9%; Score 1043.6; DB 9; Length 1077;
 Best Local Similarity 97.9%; Pred. No. 2e-224;
 Matches 1049; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
 QY 438 caatggatcccgagacttttgacagcctgggattctccaccagagaagaagggactt 497
 Db 1072 CCAATGATCCGAGACTTTTGGACAGCCTTGGATTCTCCACCAGAGAAGAGGGGACTT 1013

QY 515 tccatgaggagcatttggggcagatacagagatatggaatcagattattcaggaagg 574
 Db 481 TCCAGTGGAGGCATTTTGGGCGCAATACAGAGATATGAATCAGATTATTCAGACAGG 540
 QY 575 gaggtagcaactgcaagagtgattgacacccatcaaaacccacccctgacagacagaagaa 634
 Db 541 GAGTTGACCAACTGCAAGAGTGATGACACCATCAAAACCCACCCCTGACGACAGAAGAA 600
 QY 635 tcatcatggtggtggaatccaagagattctcctctgatgagcgtcctccatgcagtg 694
 Db 601 TCATCATGTGCGCTTGGAAUCCAGAGATCTTCCCTCTGATGGCGCTGCCCTCCATGCCATG 660
 QY 695 cccctgtccagttctatgtgtggaacagtgagctgtcctgtccagctgtaccagagatcgg 754
 Db 661 CCTCTGCCAGTTCTATGTGTGAACAGTGAGTGCTCTGCGCAGCTGTACACAGATCGG 720
 QY 755 gagacatgggctgtgtgtgcttcaacatgcagcagtgagcctgtcctcctcactacatga 814
 Db 721 GAGACATGGGCTGCGTGTCTTCAACATGCCAGCTAGCGCCCTGCTCAGCTACATGA 780
 QY 815 ttgcgcacatacagggcctaagccaggtgactttatatacacactttgggagatgcacata 874
 Db 781 TTGCGCACATACGGGCTTGAGCCAGGTGACCTTATACACACTTTGGGAGATGCACATA 840
 QY 875 ttacctgaatcacatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 934
 Db 841 TTACTGTAATCACATGAGCCACTGAAATTCAGCTTACGAGAGAAACCCAGACCTTTCC 900
 QY 935 caaagctcaggtattctgaaagtgagaaaatgatgaacttcaagagtgaaagactttc 994
 Db 901 CAAAGCTCAGGATCTTCGAAAGTTGAGAAATGATGACTTCAAAGCTGAAGACTTTC 960
 QY 995 agattgaaggtacaaatcccatccactattaaatgaaatgctgtttagggtgctt 1054
 Db 961 AGATTGAAGGTACAAATCCGATCCCACTATTAAATGGAATGCTGTTTAGGGTGCCT 1020

RESULT 5
 AL562750/c 1021 bp mRNA linear EST 16-FEB-2001
 LOCUS AL562750 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC018YA14 3
 DEFINITION prime, mRNA sequence.
 ACCSSION AL562750
 VERSION AL562750.1 GI:12911478
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1021)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT Genoscope
 Genoscope - Centre National de Sequencage
 Bp 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 I. :1021
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC018YA14"
 /clone_lib="LTI_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact : Feng liang life
 Technologies, a division of Invitrogen 9800 Medical Center

FEATURES

source

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
 8371 Email : lifang@lifetech.com URL :
 http://fulllength.invitrogen.com*
 BASE COUNT 276 a 230 c 203 g 307 t 5 others
 ORIGIN
 Query Match 61.5%; Score 945.2; DB 9; Length 1021;
 Best Local Similarity 98.0%; Pred. No. 2.9e-202;
 Matches 964; Conservative 3; Mismatches 16; Indels 1; Gaps 1;
 QY 529 ttggggcagaatacagagatatggaatcagattattcagacagggagttgaccaactg 588
 Db 984 TTTGGGGCGAGAAAACAGAGATATGAATCAGATTATTGAGAACAGGGAGTTGACCAACTG 925
 QY 589 caaagagtattgacacattcaaaacccctgcagcacagaagaatacatcatgtcgct 648
 Db 924 CAAAGAGTGTATMACACCATCAAAACCCCTACACAGAGAATCATGTGCGCT 865
 QY 649 tggaaat-ccaaagagattctcctgtgatggcgctgcctccatgcctcctgtgccagtt 707
 Db 864 TGGAAATCCCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTT 805
 QY 708 ctatgtgtgaaacagtgaagctgtcctgcagctgtaccagagatcgagagacatgggct 767
 Db 804 CTATGTGTGTAACAGTAGCTGTCTCTGCCAGCTGTACAGAGATCGGGAAACATGGGCT 745
 QY 768 cgtgtgcctttcaacatgcagctacgcctgtcctgcctgcctacatgattgcgcacatcac 827
 Db 744 CGGTGTGCTTTTCAACATCGCCCACTACGCCCTGTCTACGTACATGATTGGCACATCAC 685
 QY 828 gggcctgaagccaggtgactttatacacactttgggagatgcacatatattacotgaatca 887
 Db 684 GGGCTGAAGCCAGGTGACTTTATACACACTTTGGGAGATGCACATATTTACCTAAATCA 625
 QY 888 catgagccactgaaatcagcttcagcgagaaacccagaccttcccaagctcagagat 947
 Db 624 CATCGAGCCACTGAAATTCAGCTTCAGCGAGAAACCCAGACCTTTCCCAAGCTCAGGAT 565
 QY 948 tctctaaagtgagaaaatgatgacttcaagctgaagcttcaagattgaagggtta 1007
 Db 564 TCTTCGAAAGTTGAGAAAATGATGACTTCAAGCTGAAGCTTCAGATTGAAGGGTA 505
 QY 1008 caatccgcataccaaactattaaaatggaaatggctgttaggtgcttcaaaaggagctg 1067
 Db 504 CAATCCGATCCAACTATTAAATGGAAATGGCTGTTTAGGGTGCTTTCAAAGGAGCTCG 445
 QY 1068 aaggatattgtcagttcttaggtgtgggtgtggtgagtcgcagtgaaagcttctttgtctc 1127
 Db 444 AAGGATATTGTGAGTCTTTTAGGGTTGGCTGGATCCCGAGGTAAAGTCTTTTGTGCTC 385
 QY 1128 taaaagaagaagaaactaggtcaaaaatctgtccgtgacctatcagttattatttttaa 1187
 Db 384 TAAAGAAAAGAAAGAACTAGGTCAAAAATCTGTCCGTGACCTATCATGTTATTAATTTTAA 325
 QY 1188 ggaattgtccactggcaaatgtaactgtccagttcttccataataaaaaggctttagt 1247
 Db 324 GGATGTTGCCACTGGCAATGTAACGTGCGAGTCTTTCCATAATAAAGGCTTTGAGT 265
 QY 1248 taactcactgaggttatctgcacatgtctgaggttatgaacaaagtggaggaatgaagt 1307
 Db 264 TAACACTGAGGGTATCTGCAATGCTGAGGTATATGAACAAGGTGAGGGAATGAAGT 205
 QY 1308 tagtgtcttagcaaaaacatgtatgtcatttcaatccacgtactataaaagaagt 1367
 Db 204 TATGTGCTCTTAGCAAAAACATGATGTGATTTCAATCCCACGTACTTATAAAGAGGT 145
 QY 1368 tgggtattccaagctatttttgggaattttttagaatttttaagaattttcaaacg 1427
 Db 144 TGGTGAATTTACAGAGCTATTTTGGAAATATTTTGAATAATTTTAAGAAATTTCAACAGC 85
 QY 1428 tattccctcaaatctgagggagctgagtaaacacatcgatcatgatgtagtggtggtta 1487


```

RESULT 15
AL578646/c
LOCUS AL578646 LFL_NFL006_PL2 Homo sapiens cDNA clone CS0DK004YL08 3
DEFINITION AL578646 LFL_NFL006_PL2 Homo sapiens cDNA clone CS0DK004YL08 3
ACCESSION AL578646
VERSION AL578646.1 GI:12942922
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK004YL08"
/clone_lib="LFL_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 256 a 200 c 183 g 264 t 3 others
ORIGIN
Query Match 57.8%; Score 887.4; DB 9; Length 906;
Best Local Similarity 99.1%; Pred. No. 2.8e-189;
Matches 899; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY 605 caatcaaaacacccctgacgacagaagaatcatcatgtgcgttggaatccaagagatc 664
DB 906 CCATCAAAACCAACCCCTGACACAGAGAATCATCATGTGCGCTTGGATCCAGAGATC 847
QY 665 ttccctgtatggcgctgctccatgccatgccctgtccagttcttatgtggtgaacagt 724
DB 846 TTCCTCTGATGGGCGTGCCTCCATGCCATGCCCTCTGCCAGTTCTATGT-GTGAACAGTG 788
QY 725 agctgtctccagctgtaccagacatcgagacatggccctcgctgtgctttcaaca 784
DB 787 AGCTGTCTGCCAGCTGTACCAGATCGGAGACATGGGCCCTCGGTGTGCTTTCRACA 728
QY 785 tcgcccagctacgcccctgctcaogctacatgattggcaccatcacgggcctgaagccagtg 844
DB 727 TCGCCCACTAGCCCTGCTCAGTACATGATGGCCATCATCGGGGCGCTGAAGCCAGTG 668
QY 845 accttatacacactttggagatgcacattattaccctgaatcacatcgagccactga 904
DB 667 ACTTTATACACACTTTGGGAGATGCACATATTACCTGAATCACATCGAGCCACTGAAA 608
QY 905 ttcagcttcagcgagaaacccagactttccaaagctcaggattcttcgaaaagtga 964
DB 607 TTCAGCTTCAGCGAGAACCCAGACCTTCCCAAAGCTCAGGATTCCTCGAAAGTTGAGA 548
QY 965 aaattgatgacttcaagctgaagactttcagattgaagggtgacaaatccgcatocaa 1024

```

```

DB 547 AAATTGATGACTTCAAAGCTGAAGACTTTTCAGATTGAAGGTTACATCCGATCCAACTA 488
QY 1025 ttaaatggaaatggctgttttaggtgctttcaaaagagctcgaagatattgtcagttc 1084
DB 487 TTAAATGGAATGGCTGTTTAGGGTGCTTTCAAAGGAGCTTGAAGGATATTGTTCAGTCT 428
QY 1085 ttaggggtgggctgagatcccgagtgaaagtctcttttctctaaagaaagaaggaact 1144
DB 427 TTAGGGGTGGGCTGGATGCCGAGGTAAAGTCTCTTTTCTCTAAAGAAAGGAAGAACT 368
QY 1145 aggtcaaaatctgtccgtgacctatcagttattatttttaaggatgttgccactggca 1204
DB 367 AGGTCAAAATCTGTCGCGAGACTATCAGTTATTATTTTAAAGGATGTGCCACTGGCA 308
QY 1205 aatgtaactgtgccagttctttccataataaaaggctttgagtttaactcaactgaggggtat 1264
DB 307 AATGTAATGTGCCAGTCTCTTTCCATATATAAAAGGCTTTGAGTTAACTCACTGAGGTAT 248
QY 1265 ctgacaatgctgaggttatgaacaaagtgaaggaaatgaatgtatgtctcttagcaaa 1324
DB 247 CTGCAATGTGAGGTTATGAACAAAGTGAGGAGATGAATGTATGTCTCTTAGCAAA 188
QY 1325 aacatgtatgtgatttcaatccacgtacttataaagaagggttgatgaattcacaaagc 1384
DB 187 AACATGATGTGCATTTCAATCCCGACTTATAAAGAGGTTGGTGAAATTCACAAGC 128
QY 1385 tattttgggaatatttttagaataatttaagaatttcacaagctattccctcaaatctga 1444
DB 127 TATTTTGGATATTTTAGATATTTTAAAGATTTTCAAGCTATTCCTCAAACTGA 68
QY 1445 gggagctgagtaaacaccatcgatcatgtagatgtggttatgaactttaaaagtgtt 1504
DB 67 GGGAGCTGAGTAAACACCATCGATCATGATGATGAGTGTGTTATGAACCTTTATAGTTGT 8
QY 1505 ttatatg 1511
DB 7 TTATAWG 1

```

Search completed: July 26, 2002, 15:13:16
Job time: 26880 sec

GenCore version 4,5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2002, 09:55:26 ; Search time 5000.16 seconds
(without alignments)
6428.426 Million cell updates/sec

Title: US-09-963-333-6
 Perfect score: 1536
 Sequence: 1 ggggggggggggaccacttg.....ataataagaagtgttctgc 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

```

1: gbBa: *
2: gb_hvg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sy: *
12: gb_sts: *
13: gb_un: *
14: gb_vi: *
15: emBa: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_hvg_hum: *
31: em_hvg_inv: *
32: em_hvg_om: *
33: em_hvg_inv: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100	100	1	...
2	95	95	95	2	...
3	90	90	90	3	...
4	85	85	85	4	...
5	80	80	80	5	...
6	75	75	75	6	...
7	70	70	70	7	...
8	65	65	65	8	...
9	60	60	60	9	...
10	55	55	55	10	...
11	50	50	50	11	...
12	45	45	45	12	...
13	40	40	40	13	...
14	35	35	35	14	...
15	30	30	30	15	...
16	25	25	25	16	...
17	20	20	20	17	...
18	15	15	15	18	...
19	10	10	10	19	...
20	5	5	5	20	...

Query Match

99.7%; Score 1531.2; DB 6; Length 1536;

1	1531.2	99.7	1536	6	AR103889	Sequence
2	1531.2	99.7	1536	9	HS1SR1	Human mRNA
3	1496.2	97.4	1533	9	BC002567	Homo sapi
4	1494.8	97.3	3298	9	BC281971	Sequence
5	1484.6	96.7	1569	9	BC013919	Homo sapi
6	942	61.3	942	6	AX069365	Sequence
7	728	47.4	840	9	AB077208	Homo sapi
8	711.4	46.3	986	10	BC020139	Mus muscu
9	706.6	46.0	961	10	BC020139	Mus muscu
10	697.8	45.4	1598	10	RATHYSYN	Rattus norv
11	634.2	41.3	1349	10	MUSPSTSA	Mouse thym
12	613.2	39.9	18596	6	AR144965	Sequence
13	613.2	39.9	18596	6	AX050451	Sequence
14	613.2	39.9	18596	6	AX330682	Sequence
15	613.2	39.9	18596	6	AX330906	Sequence
16	613.2	39.9	18596	6	AX335755	Sequence
17	613.2	39.9	18596	9	HUMTS1	Homo sapien
18	613.2	39.9	43507	9	AF305037	Homo sapi
19	613.2	39.9	159974	2	AC021474	Homo sapi
20	613.2	39.9	165742	2	AC020697	Homo sapi
21	532.4	34.7	535	6	AX341901	Sequence
22	507.4	33.0	1191	5	AY005804	Danio rer
23	491.6	32.0	13610	14	AF188480	Macropodi
24	490	31.9	693	9	AB077207	Homo sapi
25	484.2	31.5	489	11	G07105	human SFS W
26	419	27.3	124138	14	AF275348	Cercopit
27	410.4	26.7	130733	14	AF017026	Macaca mu
28	408	26.6	124884	14	HEV2VXX	Varicella-z
29	406	26.4	184427	14	EHVU20824	Equine herp
30	404.6	26.3	3804	14	AF087411	Rhesus ma
31	404.6	26.3	133719	14	AF0833501	Macaca mu
32	399	26.0	10595	14	AF029302	Rhesus mo
33	396.8	25.8	108409	14	AF083424	Ateline h
34	391.8	25.5	1102	14	HSVTS4	Herpesvirus
35	391.8	25.5	43658	14	HSV3PGEN	Herpesvirus
36	391.8	25.5	112930	14	HSV2	Herpesvirus
37	390.2	25.4	3336	14	HSVTS	Herpesvirus
38	387.4	25.2	1014	14	KSU083348	Kaposi's sa
39	387.4	25.2	2528	14	AF307519	Human her
40	387.4	25.2	6600	14	KSU71365	Kaposi's sa
41	387.4	25.2	35100	6	AR065849	Sequence
42	387.4	25.2	35100	6	AR127847	Sequence
43	387.4	25.2	133661	14	U93872	Kaposi's sa
44	387.4	25.2	137508	14	KSU75698	Kaposi's sa
45	387	25.2	2428	14	AF307518	Human her

ALIGNMENTS

RESULT	1				
AR103889		AR103889	1536 bp	DNA	linear
LOCUS					
DEFINITION		Sequence 1	from patent	US 6087489,	

BASE COUNT	390 a	369 c	399 g	378 t
ORIGIN				

99.7%; Score 1531.2; DB 6; Length 1536;


```

/db_xref="taxon:10090"
/clone="MGC:28245 IMAGE:3994204"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMTV."
/clone_lib="NCI_CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
50...973
CDS
/ocdon_start=1
/product="Unknown (protein for MGC:28246)"
/protein_id="AAH20139.1"
/db_xref="GI:18044897"
/translation="MLVYSELQSDAQLSAEAPHOGELOVLRVHILRCGPKKEDR
TGNLSVFGMAARYSLRDEPPLITTRVFWKGVLELLWFKSTNAKELSSKGVRI
WDANGSRFLDSIGFSARQEDLGPVGFQWRHFGAEYKDMDSYSGQVDQLQKVID
TINPDDRRITMCANPKDPLMALPFCALCOFYVYNGELSCOLYORSQDMLGVP
FNIASVALLTYMIAHTIGLQDFVHTLGDALHLYNHLIEPLKIQLOREPRFPFKLIL
RKVEIDDFKVEDFQIEGYNPHPTIKMEMAV"
BASE COUNT 270 a 226 c 253 g 237 t
ORIGIN
Query Match 46.3%; Score 711.4; DB 10; Length 986;
Best Local Similarity 85.8%; Pred. No. 3.5e-137;
Matches 790; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 126 gctgcgcgcgcgccttgcctcccccgcgcacagcagcgcgcgcgcgcgcgcgcgcgc 185
DB 52 GCTGCTGTTGGCTCCGAGCTGCAGTCCGATGCTCAGCAGCTGAGCGCGGAAGCCGCA 111
QY 186 gcacggagctgcagtcactcctggcagatccacacatcctcgcgtcggcgtcagaaa 245
DB 112 GCATGGAGAACTCCAGTACCTGAGGAGGTGGACACATTTTGGCTGGCGGTTCAAGAA 171
QY 246 gcacgcgcgcgcgcgcgcgcgcctcgtgattcgcgcgcgcgcgcgcgcgcgcgc 305
DB 172 GGAGACCCCGACGGCAGCAGCCCTGTCGGTGTTCGGACATGCAGCGCAGATCAGCT 231
QY 306 gagatgaattccctcgtcgcacacacacacacacacacacacacacacacacacac 365
DB 232 GAGAGATGAATTCCTCTGCTCACAACCAACAGAGTGTCTGGAAGGGTGTTCGAGGA 291
QY 366 gttgctgtggtttataaaggatccacaaatgctaagagctgtcttccaaaggagtgaa 425
DB 292 GTTGTGTGTTTATCAAGGGATCCAAATGCTAAAGAAATTTGCTCCAAAGGAGTGAG 351
QY 426 aatctggatccaatgattccgcagacttttgacagcctggatcttccaccagaga 485
DB 352 AATCTGGGATCCAAATGGATCCGACAAATTTCTGGACAGCTGGGATTTCTGCCGACA 411
QY 486 agaagggagcttggcccgactttatggctccagtgaggcgcatttggggcagaatacag 545
DB 412 GGAAGGGACCTGGGCCAGCTTTATGTTTCCAAATGAGGCAATTTTGGAGCAGTACAA 471
QY 546 agatatgaatcagattattcagcagcaggaggtgaccacactgcacagagatgaacac 605
DB 472 AGATATGATTCAGATTACTCGGGAGCAAGAGTAGACCAAGCTGCAAAAAGTGTGACAC 531
QY 606 catcaaaacacacccctgacacagaagaatcatatgctgccttggaatccaaagatct 665
DB 532 CATCAAAACCAACCCCTGATGACAGAGAAATCATCATGTGCTGCTGGAAACCAAAAGAT 591
QY 666 tctctgatggcgtgctcctccatgccctcctcgcagttcttatgtgtgaacagtga 725
DB 592 TCCCTGATGGCACTGCTCCTTGGCAATGCCCTCTGTCAGTTCTATGTGTGAATGGGA 651
QY 726 gctgtcctgcagctgtaccagagatcggggagacatgggacctcgtgtgcttcttaaat 785
DB 652 ACTGTCTTGCCAGCTTTACAGAGAGTTCAGGAGATATGGGTCTGGCGCTGCCCTTCAACAT 711
QY 786 cgcacgactacgcctcgtcactcagttacatgattgcacatacagcgcgcgcgcgcgc 845

```

```

DB 712 TGCCAGCTATGCTCTGCTACCTACATGATTGCACATATCACAGGCTGCAGCCAGGTGA 771
QY 846 cttatcacactcttgagatgcacatatatttaactgaatcacatcagccactgaaaat 905
DB 772 TTTTCTCCACACTTTGGGAGATGCACATATTACTGAATCATATAGAGCCGCTGAAAT 831
QY 906 tcagcttcagcagacaccagcacttctcccaagctcagagattcttcgaaaagttagaaa 965
DB 832 TCAGCTACAGCAGAGAACCAAGACCTTTCCTCAAGCTCAAAAGTTCGAAAAGTTGAGAC 891
QY 966 aattatgaacttaaacagcagacttccagattgaagggtacacatcccgactcaactat 1025
DB 892 AATCGATGATTCAAGATTGGAAGACTTTTCAGATGAAGGGGTATATATCCACATCCACCAT 951
QY 1026 taaaatgaaatgctgttta 1046
DB 952 TAAATGGAAATGGCTCTTTA 972
RESULT 9
MUSTHSN 961 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse thymidylate synthase mRNA, complete cds.
DEFINITION M13019
ACCESSION M13019.1 GI:202029
VERSION thymidylate synthetase.
KEYWORDS Mouse, cDNA to mRNA, clone pMTS-3.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 961)
AUTHORS Petryman, S.M., Rossana, C., Deng, T., Vanin, E.F. and Johnson, L.F.
TITLE Sequence of a cDNA for mouse thymidylate synthase reveals striking
similarity with the prokaryotic enzyme
JOURNAL Mol. Biol. Evol. 5, 313-321 (1988)
MEDLINE 88174353
COMMENT Draft entry and printed copy of the sequence [1] kindly provided by
L.F. Johnson, 29-AUG-1986.
FEATURES
source Location/Qualifiers
1..961
/organism="Mus musculus"
/db_xref="taxon:10090"
<1..961
/product="ths mRNA"
38..961
/notes="thymidylate synthase (EC 2.1.1.45)"
/codon_start=1
/db_xref="GI:202030"
/translation="MLVYSELQSDAQLSAEAPRHGELOVLRVHILRCGPKKEDR
TGTGTLVFGMAARYSLRDEPPLITTRVFWKGVLELLWFKSTNAKELSSKGVRI
WDANGSRFLDSIGFSARQEDLGPVGFQWRHFGAEYKDMDSYSGQVDQLQKVID
TINPDDRRITMCANPKDPLMALPFCALCOFYVYNGELSCOLYORSQDMLGVP
FNIASVALLTYMIAHTIGLQDFVHTLGDALHLYNHLIEPLKIQLOREPRFPFKLIL
RKVEIDDFKVEDFQIEGYNPHPTIKMEMAV"
BASE COUNT 253 a 224 c 251 g 233 t
ORIGIN 63 bp upstream of PstI site.
Query Match 46.0%; Score 706.6; DB 10; Length 961;
Best Local Similarity 85.5%; Pred. No. 3.4e-136;
Matches 787; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 126 gctgcgcgcgcgccttgcctcccccgcgcacagcagcgcgcgcgcgcgcgcgcgcgc 185
DB 40 GCTGCTGTTGGCTCCGAGCTGCAGTCCGATGCTCAGCAGCTGAGCGCGGAAGCCGACG 99
QY 186 gcacgggagctgcagtcactcctgggagatccacacatccctcgcgtcgcgcgcagaaa 245
DB 100 GCATGGAGAACTCCAGTACCTGAGCAGGAGTGGAGCACAATTTTGGCGCTGCCGCTTCAAGAA 159
QY 246 gcacgcgcgcgcgcgcgcgcgcctcgtcgtattcgcgcgcgcgcgcgcgcgcgcgc 305

```

Db	524	CGACAGGAAGGGGACCTGTGGGCCCAAGTTATATGGTTTCCAAATGGAGGAGCAATTTTGGACCGAGAG	583
QY	541	tacagagatggaatcaagattattacagacagcagggagttgaaccaatcgcaaaagagtatt	600
Db	584	TCAAAGACATGGATTCAAGATTACTCAGGCAACAGGAGTAGACCAGCTGCARAAAGTGATT	643
QY	601	gacacatcaaaacaaacacctgacacacagagaagaatcatcatgtgpgcttggaaatccaaga	660
Db	644	GACACCATCAAAACCAACCTGATGACAGAAGANTCATATGTGTGCTTGGAAACCCAAA	703
QY	661	gatottctctgatggg-----ctgcctcaatgcgatgcctctgccagttctatgtggt	716
Db	704	GATCTTCCCTGTATGTCGCACTGCCTGCCCTCTGGCATGCCCTCTGTCAATTCTATATGGT	763
QY	717	gaacagtgatgctcctgccagctctacacagagatcggagacatcgggacctcgtgtgccc	776
Db	764	GAATGGGAGGCTATCTTCCAGCTTTACACAGAGTCAAGAGATATGGGTCTGGGTGTGCC	823
QY	777	tttaacatcgccagtcagcctgctcacgtacatgattggcgacatcagggcctg--	834
Db	824	TTTCAACATTGCCAGCTATGCTCTGCTCACCTACATGATTGGACATATCACAGGCGCTGCA	883
QY	835	-----aagcagtgagctattatacacactttgggagatgcacatatttacctgaatc	886
Db	884	GACATTTTATGCCAGGTGATTTTGTCCCAATTTGGGAGATGCACATATTTACCTGAATC	943
QY	887	acatcgacacactgaaaattcagcttcacgcgagaaacccagacaccttcccaagctcagga	946
Db	944	ATATAGACCCCTGAAATTCAGCTACAGCG-----AGACCTTTGCCCAAMAGCTCAAAA	996
QY	947	ticttcgaasgtgagaaaattgatgacttcaagctgaagacttcagattgaagggt	1006
Db	997	TCTTTGSAAGAGGTGAGCAATCGATGATTTCAAAGTTGAAGCTTTCAGATTGAGGGT	1056
QY	1007	acaatccgcacccaactattaaaatggaaatggctgtttaggtgcttccaaaggagctc	1066
Db	1057	ATATCCACATCCACCATTAATAATGGAGATAGCTGTTTAGAGAGCTTTACGTGATGCTG	1116
QY	1067	gaagat 1073	
Db	1117	TGAATAT 1123	
RESULT 12			
ARI144965			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Query			
Db			

```
JOURNAL      gene sets
Patent: WO 0194629-A 1191 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
  source      1..18596
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
BASE COUNT   4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match      39.9%; Score 613.2; DB 6; Length 18596;
Best Local Similarity 90.6%; Pred. No. 9.8e-117;
Matches 667; Conservative 0; Mismatches 63; Indels 6; Gaps 1;
QY 807 gtacatgatgcgcacatcaacgagcctgaagccaggtgactttatcacacactttgggaga 866
Db 15510 GAACCTTTGTTGATCACAATCTGTTACTTGTTCACGGACATGAGGACCAATTACAACAG 15569
QY 867 tgcacattatcaactgaatcacatcagccacagcaaaattcagcttcagcgagaaaccag 926
Db 15570 GTCGTACAAATATGCGAAATAATGCGCTTATTTTTCAGGACATGAGGACCAATTACAACAG 15629
QY 927 acccttccaaagctcaggaattcttgcgaaggttgagaaattgatgacttcaaaagtga 986
Db 15630 ACCTTTCCAAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTTGATCACTTCAAAGCTGA 15689
QY 987 agactttcagattgaaggtacaatccgcacatcccaactattaaatggaaatggctgttta 1046
Db 15690 AGACTTTTCAGATTGAGGGTACAAATCCGCATCCCACTATTAAATGGAATGGCTGTGTTA 15749
QY 1047 ggggtcttcaaaagagctcgaagatatgtcagctctttaggggttggctggatgccg 1106
Db 15750 GGGTCTTTCAAAGAGCTCGAAGGATTTGTCACTCTTTAGGGGTGGGCTGGATGCCG 15809
QY 1107 aggtaaaagttcttttgcctcaaaagaaaggaactaggtcaaaaactctgcccgtgac 1166
Db 15810 AGGTAAAAGTTCTTTTGGCTCTAAAAGAAAAGAGAACTAGGTCAAAAATCTGTCGTGAC 15869
QY 1167 ctatcagttatttaatttttaaggtatgtgccactggcgaatgttaactgtgcagttcttt 1226
Db 15870 CTATCAGTTATTAAATTTTAAAGGATGTGCCACTGGCAATGTAACTGTGCCAGTTCITT 15929
QY 1227 ccataataaaggcttgagttgaactcaactgaggttatctgacaatgctgaggttatgaa 1286
Db 15930 CCATAATAAAGGCTTGGTGAATTTTCAAGCTATTTCACAAATCTGCAATGCTGAGGTTATGAA 15989
QY 1287 caaagtggaggaatgaaatgtatgtcttagcaaaaacatgtatgtgcatattcaatc 1346
Db 15990 CAAAGTGGAGGAATGAATGTATGTGCTCTTAGCAAAAACATGTATGTGCATTTCAATC 16049
QY 1347 ccactactataaaggcttggtgaattccaaagctatttttggaaatatttttagaa 1406
Db 16050 CCACGTACTTATAAAGAGGCTGGTGAATTTTCAAGCTATTTCACAAATCTGCAATGCTGAGGTTATGAA 16109
QY 1407 tattttaagaatttcaaacgtattccctcaaatctgaggagctgagtaaacaccatcga 1466
Db 16110 TATTTTAAGAATTTCAACAGTATTCCCTCAAACTCTGAGGAGCTGAGTAACACCATCGA 16169
QY 1467 tcatgatgaggtgtgtttatgaacttttaa-----agttgttttatatgttgctataa 1520
Db 16170 TCATGATGTAGANGTGTGGTTATGAATTTAAAGTTATAGTTGTTATATGTGCTATAA 16229
QY 1521 taagaagtgttctgc 1536
Db 16230 TAAAGAAGTGTCTGC 16245
RESULT 15
AX330906
LOCUS
DEFINITION Sequence 1415 from Patent WO0194629.
ACCESSION AX330906
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2002, 17:46:36 ; Search time 77.68 Seconds
(without alignments)
3753.437 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgactcactca.....gagggagcgcgctgggga 1187

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1163.8	98.0	18596	4	US-09-318-448-11
2	297.8	25.1	1536	3	US-09-089-195-1
3	191.8	16.2	14747	4	US-09-608-285A-42
4	191.8	16.2	15977	4	US-09-608-285A-59
5	189.4	16.0	3804	2	US-08-483-488-5
6	189	15.9	631	4	US-09-385-982-354
7	189	15.9	3867	4	US-09-347-114A-81
8	189	15.9	35060	3	US-08-814-095-7
9	188.8	15.9	2713	2	US-08-916-901-6
10	188.8	15.9	2713	4	US-09-154-602-6
11	188.8	15.9	87350	3	US-08-781-891-79
12	187.6	15.8	17327	1	US-07-906-871-15
13	187.2	15.8	72604	4	US-09-268-992-7
14	186	15.7	8453	1	US-09-167-681-45
15	185.8	15.7	6769	1	US-08-480-784-20
16	185.8	15.7	6769	1	US-08-483-553-20
17	185.8	15.7	6769	1	US-08-487-002-20
18	185.8	15.7	6769	1	US-08-483-554B-20
19	185.8	15.7	6769	1	US-08-488-011B-20
20	185.8	15.7	6769	4	US-08-850-727-20
21	185.8	15.7	6769	5	PCT-US95-10202-20
22	185.8	15.7	6769	5	PCT-US95-10203-20
23	185.8	15.7	6769	5	PCT-US95-10220-20
24	183.6	15.5	619	4	US-09-385-982-358
25	183.4	15.5	84495	4	US-09-797-906-3
26	182.6	15.4	11288	4	US-08-646-301A-1
27	182.6	15.4	11288	4	US-08-481-968A-4

c	28	182.6	15.4	11288	4	US-08-154-712B-4	Sequence 4, Appli
c	29	182	15.3	59065	4	US-09-813-817-3	Sequence 3, Appli
c	30	181.6	15.3	87350	3	US-08-781-891-79	Sequence 79, Appli
c	31	181.4	15.3	4038	3	US-08-969-125-8	Sequence 8, Appli
c	32	180.8	15.2	1811	1	US-08-848-252-1	Sequence 1, Appli
c	33	179.4	15.1	841	5	PCT-US93-06251-80	Sequence 80, Appli
c	34	179.4	15.1	841	5	PCT-US93-06251-81	Sequence 81, Appli
c	35	179.4	15.1	1386	2	US-08-687-080-76	Sequence 76, Appli
c	36	179.2	15.1	5835	4	US-09-033-333-3	Sequence 3, Appli
c	37	179.2	15.1	5836	1	US-08-380-916-1	Sequence 1, Appli
c	38	179.2	15.1	5836	3	US-08-721-690-1	Sequence 1, Appli
c	39	179.2	15.1	5836	3	US-08-891-581-1	Sequence 2, Appli
c	40	179.2	15.1	5836	4	US-09-033-333-2	Sequence 11, Appli
c	41	178.6	15.0	1988	2	US-08-257-963A-11	Sequence 11, Appli
c	42	178.6	15.0	1988	4	US-08-367-841A-11	Sequence 11, Appli
c	43	178.6	15.0	1988	5	PCT-US95-07201-11	Sequence 11, Appli
c	44	178.6	15.0	2562	2	US-08-436-771-8	Sequence 8, Appli
c	45	178.6	15.0	2562	2	US-08-434-998-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-318-448-11
; Sequence 11, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-11

Query Match	98.0%	Score	1163.8	DB	4	Length	18596
Best Local Similarity	99.1%	Pred. No.	4.3e-194				
Matches	1176	Conservative	0	Mismatches	10	Indels	1
QY	1	gatcgccactgactcactcagcctgggtgagagagagagactctgtctcaaaaaa	60				
Db	78	gatcgccactgactcactcagcctgggtgagagagagagactctgtctcaaaaaa	137				
QY	61	aaaaagcccgagggctcaaaaaaaaccccgaaagcccgctgggtgttttttt	120				
Db	138	aaaaagcccgagggctcaaaaaaaaccccgaaagcccgctgggtgttttttt	197				
QY	121	tt	180				
Db	198	tt	257				
QY	181	cggtattgtgctcactgcaacccctgctccaggtttcaagcaattcttctgctcagcc	240				
Db	258	cggtattgtgctcactgcaacccctgctccaggtttcaagcaattcttctgctcagcc	317				
QY	241	tcccaagtaccaccacgcccagctaattttttttttttttttttttttttttttt	300				
Db	318	tcccaagtaccaccacgcccagctaattttttttttttttttttttttttttttt	377				
QY	301	catgtttcagctggtgtgtaactcctgactcaggtgatcccccgcctcgccccc	360				
Db	378	catgtttcagctggtgtgtaactcctgactcaggtgatcccccgcctcgccccc	437				
QY	361	caaatgtactaggattacagggcgtgagccacggctccagcgccgtgggtttttaaca	420				

[illegible]

; NAME/KEY: exon

NAME/KEY: exon

Search completed: July 26, 2002, 17:48:22
Job time: 12551 sec

Db	1338	GGCANGATCTCGGCTACTGCAAACCTCTGGCTCCCGGGTTCAAGTGATTCCTCTGCCACA	1299
QY	238	gcctccccaagta-----gccacacgcccagcgtaa- <u>ttt</u> tgtan	276
Db	1298	GCCCTCTGAGTAGCTGGGATTACAGSCATGGCGCACCGCGCCAGCTAA <u>TTTTTTT</u> TGTAT	1239
QY	277	tttttagtagagacgggggtttcaccaatgttgtccaggctgggtctgaactcotgaccctca	336
Db	1238	TTTTTAGTAGAGACAGGGGTTCACCATGTTGGCAGAGCTGGTCTCGAACCTCTCGACCTCA	1179
QY	337	ggtgatccaccgcgcctcgcccccaagtactaggattacaggcgtgagccaccgcgc	396
Db	1178	GGTGATCCACCGCCTCGGCCCTCCCAAAGTCTGGGATTACAGGCATGAGCCACCGGCC	1119
QY	397	cagc	400
Db	1118	TTGGC	1115

```

RESULT 11
US-08-781-891-79
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-79

```

Query Match 15.9%; Score 188.8; DB 3; Length 87350;
Best local Similarity 69.9%; Pred. No. 6.6e-25;
Matches 320; Conservative 0; Mismatches 108; Indels 30; Gaps 4;

QY 1 gatcgccactgcactccagcctgggtgagagcagcactctctcaaaaaaaaaa 60
DB 29750 GATCGGCCACTGCATCTCCAGCCTGGGTGACAGCAGNNNACTCCGTC 29809
QY 61 aaagaacgccgggctcaaacaaaaaacctcgaaaaccctggcggtctttttt 120

```

Db 29810 AAAAATAAAATAAATAATCTCAGACCATACATTTAATGAGGTGTTTTTGGTTTTT 29860
QY 121 tttttttttttttttttt-----gggaagctctgtctgtcgcacagctgagtt 172
Db 29870 TTTTGTGTTTTTTTTTTTTTTTGGTGATAGAGTTCTCACTCTCTCACCTAGCTAGAGT 29929
QY 173 acaatggtcgaatctt--ggtcactcaaacctctgctccacaggttcaagcaatcttc 230
Db 29930 GCAGTGGCGGGAATCFCNGGCTCACTGCAACCTCCGCCCTCCGGTGTCAAGCATTCTCC 29989
QY 231 tgctcagcgcctcccaagta-----gccacacacgccagcgaatttt 271
Db 29990 TGCTTCAGCTCTCGGAATAGCTGGAGCTACAGTGGCGGTGCCACCCCGGCTAATTTT 30049
QY 272 tytanttttagtagaagcgggggttccaatgttgtccagctggtctcngaactctga 331
Db 30050 TTGTATTTAGTAGAGA-TGAGGTTTCACTGTGTTGTCACAGSCTGGTGTTTGAACCTCTGA 30108
QY 332 cctcagtgatcacccgcctcggcccccacaaagtactagattacaggcgtgagccacc 391
Db 30109 GGTCAAGCAATACACCCGCCCGGCGCTCCCAAAATGTTGGGATACAGGCGTGAGCCACA 30168
QY 392 ggcctcagcgcctcggcggttttttaatcaagtagaaaa 429
Db 30169 GTGCTGGCCAGAGAGATATTTAATGAAAAATAATA 30206

RESULT 12
US-07-906-871-15/C
; Sequence 15, Application US/07906871
; Patent No. 5340739
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Avraham, Shalom
; TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN A
; NUMBER OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1255 Connecticut Avenue, N.W., Suite 300
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,871
; FILING DATE: 19920103
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/816,289
; FILING DATE: 03 JAN 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/635,544
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US89/03051
; FILING DATE: 13-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/224,035
; FILING DATE: 13-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimballa, Michele A
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0627.2830004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)833-7533

```


OY 1521 taagaagtggtctgc 1536
|||||
Db 16230 TAAAGAAAGTGTCTGC 16245

Search completed: July 26, 2002, 16:37:26
Job time: 24120 sec

QY 867 tgacacattattacatgaatcacatcgagccactgaaattcagcttcacgagacccag 926
 Db 15570 GTCGTACAAATATGCGCAATATGCGCTTATTTTGTGTTTATGCTTCACGAGAACCCAG 15629
 QY 927 accttcccaaaagctcaggtattcttcgaaaagttgagaaaattgatgacttcaaaagtga 986
 Db 15630 ACCCTTCCCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTGATGACTTCAAAAGCTGA 15689
 QY 987 agacttcagattgaagggtacaaatccgcatcccaactattaaaaatgaaatggctgttta 1046
 Db 15690 AGACTTCAGATTATGCGCAATATGCGCTTATTTTGTGTTTATGCTTCACGAGAACCCAG 15629
 QY 1047 ggggtcttcaaaagagctcgaaggtatttgcagttctttaggggttgggctggatgccg 1106
 Db 15750 GGGTCTTTCAAAAGGAGCTCGAAGGATATTCAGTCTTTAGGGGTGGGCTGGATGCCG 15809
 QY 1107 aggttaaaagtcttttcttcttaaaagaagaagaaactaggttcaaaatctgtccgtgac 1166
 Db 15810 AGGTAAAAGTCTTTTGTCTTAAAAGAAAAGAACTAGGTCAAAAATCTGTCCTGAC 15869
 QY 1167 ctatcagttattatttttaaggtattgccaactgaggttatctgcaaatgtgaggttatgaa 1226
 Db 15870 CTATCAGTTATTAATTTTAAAGGATGTCGCACTGGCAATGTAATGTAATGTCCTTT 15929
 QY 1227 ccataataaaagctttgagtttaactcactgaggttatctgcaaatgtgaggttatgaa 1286
 Db 15930 CCATAATAAAAGGCTTTGAGTTAACTCACTGAGGATATCTGACATGCTGAGGTATGAA 15989
 QY 1287 caaagttagagagaatgaaatgtatgtctcttagcaaaacacatgtatgtgacttcaatc 1346
 Db 15990 CAAAGTGAGGAGAATGAAATGTATGTCTCTTAGCAAAACATGATGTGCAATTCATC 16049
 QY 1347 ccagctactataaaagaaggttgggaatttcacaaagctatttttgggaatttttagaa 1406
 Db 16050 CCAGCTACTTATAAAGAGGTTGGTGAATTTTCAACAGCTATTTTGGAAATTTTGTAGAA 16109
 QY 1407 tattttaagaatttcacaaagctattccctcaaatctgagggagctgagtaacacacatga 1466
 Db 16110 TATTTTAAAGATTTTCAACAGCTATTTTCCCTCAATCTGAGGAGCTGAGTACACATCGA 16169
 QY 1467 tcatgtatgagtggtgttatgaaactttaA-----agttgttttatattgttctataa 1520
 Db 16170 TCATGATGATGAGTGTGTTTAAAGTTATAGTTGTTTATATATGTTGCTATAA 16229
 QY 1521 taaagaagtgttctgc 1536
 Db 16230 TAAAGAAGTGTCTGCG 16245

RESULT 13
 LOCUS AX050451 18596 bp DNA linear PAT 12-JAN-2001
 DEFINITION Sequence 11 from Patent WO0071754.
 ACCESSION AX050451
 VERSION AX050451.1 GI:12226657
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 18596)
 Johnson, W.G. and Stearns, E.S.
 Methods for diagnosing, preventing, and treating developmental
 disorders due to a combination of genetic and environmental factors
 Patent: WO 0071754-A 11 30-NOV-2000;
 University of Medicine and Dentistry of New Jersey (US)
 FEATURES
 Location/Qualifiers
 source 1..18596
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 4521 a 3991 c 4479 g 5605 t
 ORIGIN

Query Match 39.9%; Score 613.2; DB 6; Length 18596;
 Best Local Similarity 90.6%; Pred. No. 9.8e-117;
 Matches 667; Conservative 0; Mismatches 63; Indels 6; Gaps
 QY 807 gtacatgatgcgcacatcacccggcctgaagccaggtgactttatcacacattgggaga 866
 Db 15510 GAACCTTTTGATCACATCTGTGTACTTGTTCACGACATGAGGACAAATACACAG 15569
 QY 867 tgcacatatatttaoctgaatcacatcgagccactgaaattcagcttcagcaggaacccag 926
 Db 15570 GTCGTACAAATATGCGCAATATGCGCTTATTTTGTGTTTATGCTTCACGAGAACCCAG 15629
 QY 927 accttcccaaaagctcaggtattcttcgaaaagttgagaaaattgatgacttcaaaagtga 986
 Db 15630 ACCCTTCCCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTGATGACTTCAAAAGCTGA 15689
 QY 987 agacttcagattgaagggtacaaatccgcatcccaactattaaaaatgaaatggctgttta 1046
 Db 15690 AGACTTCAGATTATGCGCAATATGCGCTTATTTTGTGTTTATGCTTCACGAGAACCCAG 15629
 QY 1047 ggggtcttcaaaagagctcgaaggtatttgcagttctttaggggttgggctggatgccg 1106
 Db 15750 GGGTCTTTCAAAAGGAGCTCGAAGGATATTCAGTCTTTAGGGGTGGGCTGGATGCCG 15809
 QY 1107 aggttaaaagtcttttcttcttaaaagaagaagaaactaggttcaaaatctgtccgtgac 1166
 Db 15810 AGGTAAAAGTCTTTTGTCTTAAAAGAAAAGAACTAGGTCAAAAATCTGTCCTGAC 15869
 QY 1167 ctatcagttattatttttaaggtattgccaactgaggttatctgcaaatgtgaggttatgaa 1226
 Db 15870 CTATCAGTTATTAATTTTAAAGGATGTCGCACTGGCAATGTAATGTAATGTCCTTT 15929
 QY 1227 ccataataaaagctttgagtttaactcactgaggttatctgcaaatgtgaggttatgaa 1286
 Db 15930 CCATAATAAAAGGCTTTGAGTTAACTCACTGAGGATATCTGACATGCTGAGGTATGAA 15989
 QY 1287 caaagttagagagaatgaaatgtatgtctcttagcaaaacacatgtatgtgacttcaatc 1346
 Db 15990 CAAAGTGAGGAGAATGAAATGTATGTCTCTTAGCAAAACATGATGTGCAATTCATC 16049
 QY 1347 ccagctactataaaagaaggttgggaatttcacaaagctatttttgggaatttttagaa 1406
 Db 16050 CCAGCTACTTATAAAGAGGTTGGTGAATTTTCAACAGCTATTTTGGAAATTTTGTAGAA 16109
 QY 1407 tattttaagaatttcacaaagctattccctcaaatctgagggagctgagtaacacacatga 1466
 Db 16110 TATTTTAAAGATTTTCAACAGCTATTTTCCCTCAATCTGAGGAGCTGAGTACACATCGA 16169
 QY 1467 tcatgtatgagtggtgttatgaaactttaA-----agttgttttatattgttctataa 1520
 Db 16170 TCATGATGATGAGTGTGTTTAAAGTTATAGTTGTTTATATATGTTGCTATAA 16229
 QY 1521 taaagaagtgttctgc 1536
 Db 16230 TAAAGAAGTGTCTGCG 16245

RESULT 14
 LOCUS AX330682 18596 bp DNA linear PAT 09-JAN-2
 DEFINITION Sequence 1191 from Patent WO0194629.
 ACCESSION AX330682
 VERSION AX330682.1 GI:18103660
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signal

```

Db 160 GGAGGACCGCAGCGGCACTGGCACCCCTGTGGTGTTCGGCATCGCAGGACGATACAGCCT 219
QY 306 gagagatgaattccctctctctcaacacaaacgctgttcttgaaggggtttttggagga 365
Db 220 GAGAGATGAATTCCTCTGCTCACACACACAGAGAGTCTTCTGAAGGGTGTTCGGAGGA 279
QY 366 gtgtgtgtgtttatcaagggttccacaaatgctaaagagctgtcttccaaaggagtgaa 425
Db 280 GTGTGTGTGTATCAAGGATCCACAAATGCTAAGAAATGTCTCAAGGAGAGTGAG 339
QY 426 aattggttgatgccaatggatcccgagacttttggacagctgggattctccacacagaa 485
Db 340 AATCTGGGATGCCAATGATCCCGGAGATTTCTGGACAGCTGGGATTTCTGCCCGACA 399
QY 486 aagagggaacttggcccaatttatggctccagtgagggcattttggggcagagaatcacg 545
Db 400 GGAAGGGAGACTGGGCCAGTTATGTTTCCAAATGAGAGCAITTTGAGCAGAGTAGTCAA 459
QY 546 agatattgaatcagattattcaggacagaggagttgacaaactgcaaaagtgattgacac 605
Db 460 AGATATGGATTAGATTAATCTCGGACAGAGTAGACCACTGCAAAAGATGATTGACAC 519
QY 606 catcaaaacaaacctgacgacagagaagaatcatcatgtgcgttggaatccaaagatct 665
Db 520 CATCAAAACCAACCCCTGATGACAGAAGAAATCATCATGTGTGCTGGAAACCAAGATCT 579
QY 666 tccctgatggcgtcctccatgccatgccctctccagttctctatgtgttgaaacagtga 725
Db 580 TCCCTGATGGCACTGCCCTCCCTGCCATGCCCTCTCTGCTAGTTCTATGTGTTGAATGGGA 639
QY 726 gctctcctcagctgtaccagagatcgggagacatcggcagatcggcgtgtgctcttcaaat 785
Db 640 ACTGCTGTGCAGCTTACAGAGGTGAGGAGATATGGGTCTGGCGCTGCCCTTCAACAT 699
QY 786 cgcagctacgcctcgtcctcagatgattgcgcacatcagcggcctgaagccagtgga 845
Db 700 TGCCAGCTATGCTCTGCTCACTACATGATGATGACATATCAAGGCTTCGAGCCAGGTGA 759
QY 846 ctttatcacacttggagatgacatatattaccctgaatcacatcgagccactgaaat 905
Db 760 TTTGTGCCACACTTGGAGATGCACATATTTACCTGAATCATATAGACCGCTGAAAT 819
QY 906 tcagcttcagcagacacacacacacacacacacacacacacacacacacacacacacacac 965
Db 820 TCAGCTACAGCAGGACCAAGACACCTTTCCCAAGCTCAAAATCCCTTGGAAAGTTGAGAC 879
QY 966 aattgatgacttcaagctgaagacttccagattggaaggttcaaatccgactccaaactat 1025
Db 880 AATCGATGATTTCAAAGTTGAAGACTTTCAGATTGAAGGTATATATCCACATCCACGAT 939
QY 1026 taaaatggaatggctgttta 1046
Db 940 TAAATGGAATGGCTGTTTA 960

```

```

RESULT 10
LOCUS RATHYSYN
DEFINITION Rattus norvegicus thymidylate synthase mRNA, linear ROD 30-MAY-2000
ACCESSION L12138
VERSION L12138.1 GI:207327
KEYWORDS thymidylate synthase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1598)
AUTHORS Ciesla,J., Weiner,K.X., Weiner,R.S., Reston,J.T., Maley,G.F. and
Maley,F.
TITLE Isolation and expression of rat thymidylate synthase cDNA:
phylogenetic comparison with human and mouse thymidylate synthases
JOURNAL Biochim. Biophys. Acta 1261 (2), 233-242 (1995)

```

```

MEDLINE 95226450 Location/Qualifiers
FEATURES
source
I. .1598
/organism="Rattus norvegicus"
/strain="Reuber"
/db_xref="taxon:10116"
/cell_line="H35"
/cell_type="hepatoma"
I. .924
/EC_number="2.1.1.45"
/function="conversion of dUMP to dTMP"
/note="dimer; putative"
/codon_start=1
/product="thymidylate synthase"
/protein_id="AAA92340.1"
/db_xref="GI:207328"
/translation="MVEGSELQSGAQOPEARPOHGELOYLRQVHEIMRCGFKI
TGTGFLSVFGMOARYSLRDEFPLLTTRKRVKMGVLELLMFKIGSTNAKELSSK
WDANGSRDFLDSLGSARQEGDLGVPYGFOWHFGADYKDMDSYSGQVDLOI
TKTNPDDRRIIMCAWNPDKDPLMALPCHALCQFYVYVNGELSCQLYQKSGDMGI
FNIAYSALLTYMTAHTGLQPGDFVHTGLDAHYLNHIEPLKIQLOREPRFPFKI
RKVETIDDKVEDFQIEGYNPHPTIKMEMAV"
BASE COUNT 438 a 328 c 375 g 457 t
ORIGIN
Query Match 45.4%; Score 697.8; DB 10; Length 1598;
Best Local Similarity 87.7%; Pred. No. 2.4e-134;
Matches 762; Conservative 0; Mismatches 107; Indels 0; Gaps
QY 181 ccgcgcacagggagctcagctacccctgggcagatccacacacatccctcgcgtggcgctc 240
Db 58 CCGCAGCATGGAGAACTCCAGTACTTCCAGTCAGCAGCAGTGGAGCAGCATATTCTGCGGGTTC 117
QY 241 aggaagacagcgcgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
Db 118 AAGAAGGAGGAGCCGACGCGGCACCTGGCACCTCTCGGTGTTCCGATCAGCAGCAGGTAC 177
QY 301 agcctgagatgaattccctcctcgtcgtgacacacacacacacacacacacacacacacac 360
Db 178 AGCCTGAGAGATGAATTCCTCTCTCTCTCACAAACCAACAGAGTATCTGGAAGGGTGTTC 237
QY 361 gagaggtcgtgtgtttatcaagggtatccacaaatgctaaagagctctctccaaagga 420
Db 238 GAGGAGTGTGTGTGTGTATTAAAGGCTCCCAAAATGCTAAAGAACTGCTCTCCCAAGGA 297
QY 421 gtcaaatctggatgccaatggatcccgagacttttggacagcctggagctctccacc 480
Db 298 GTGAGATTTGGATGCCAATGGGTCCCGAGACTTTTGGACAGCTTTTGGACAGTTCTCGCC 357
QY 481 agagaagaaggggacttggcccccagtttatggctccagtgaggcattttggggcagaa 540
Db 358 CGACAGAGAGGAGACTGGGCCAGTTTATGGATTCCAGTGGAGACATTTTGGAGCAGAC 417
QY 541 tacagagatggaatcagattattcaggacagggaggttgacaaactgcaaaagagtgatt 600
Db 418 TACAAGATGTGATTCAGATTACTCGGGTCAAGAGGATAGACCAGTTCGCAAAAGTGATT 477
QY 601 gacacatacaaaccaacccctgcgcagacagaagaatcatcgtgcgtctggaatcccaaga 660
Db 478 GACACATCAAAACCAACCCCGATGACAGAAAGTAATCATCATGTGTCGCTGGAACCCAAA 537
QY 661 gatcttcctctgatggcgtcctccatgccatgcctctgcagattctctatgttgtgaac 720
Db 538 GATCTTCCCTGATGGCACTGCCCTCTTGGCATGCCCTCTGTCATATTATATGTGGTGAAT 597
QY 721 agtgagctcctcgcagcgtgtaccagagatcgggagacatgggctcgtgtgctcttc 780
Db 598 GGGGAGCTCTCTGCGAGCTTTACACGGGTACAGGAGATATGGGTCTGGGTGTGGCTTC 657
QY 781 aacatgcagctacgcctcgtcagcagatgattgagcacaacagggcctgaagaca 840
Db 658 AACATGGCAGCTATGCTCTGCTGACCTACATGATTCACATATACACGGGCTTCGACCG 717

```

QY 586 CTGCAAGAGTgattgacacccatcaaaaccaaccctgacgacgagaagaatcatctatgtgc 645

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 36 Row: j Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
 Location/Qualifiers
 1..986
 /organism="Mus musculus"

[illegible][illegible]

Db 541 GGGAGTGTGACAACTGCAAAAGAGTGATTGACACACATCAAAACCAACCTGTGACGACAGAG 600

QY 633 aatcatcatgtgccttggaatccaagagatcttccctctgatggcgctgcctccatgcaa 692

Db 601 AATCATCATGTGGCTTGAATCCAGAGATCTCTCTGTATGGCGCTGCCTCATGCCA 660

QY 693 tgcctctgcagttctatgtgtgaacagtgagctgtcctctccagctgtaccagatgc 752

Db 661 TGCCCTCTGCCAGTTCATGTGGTGACACAGTGAGCTGTCTGCGCAGCTGTACCAGATC 720

QY 753 gggagacaatgggctcgtgtgtcctttcaacatcgcagctcgcctcgtccatgtacat 812

Db 721 GGGAGACATGGGCTCGGTGTCCTTTCAACATGCGCAGCTGACGCCCTGCTACAGTACAT 780

QY 813 gattgcgcacatcacggcgctgaagccaggtgacctttatcacacac--tttgggagatgca 870

Db 791 GATTGCCACATCACGGGCTGAAGCCAGGTGACTTTATACACACCTTTGGGAGATGCA 840

QY 871 catattacacgaatcacatcagccactg--aaaattcagcttcagcg--agaaccaga 927

Db 841 CATATTACCTGAATCCATCGAGCCACTTGGAATTTTCAGCTTCAGCGAAGAACCCAGA 900

QY 928 ccttt-cccaactcagattctt---cgaaaagttgagaaaattgatga-cttcaag 982

Db 901 CCTTCCCAAGCTCAGGATTCCTTCCAAAAGGTTGAGAAAATTGATGACCTTCCAG 960

QY 983 ctgaagacttt--cagattgaaggtacacatccqcatccaactattaaaatggaaatggc 1040

Db 961 CTGAAAATTTTCAGATTGAGGGACAACTCCGCTCATCTTTTAAATGGAAATGGG 1020

QY 1041 t 1041

Db 1021 T 1021

RESULT 14

LOCUS AL515910/c

DEFINITION AL515910 LFI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA001YD20 3

prime, mRNA sequence.

ACCESSION AL515910

VERSION AL515910.1 GI:12779403

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 923

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DA001YD20"

/clone_lib="LFI_NFL011_NBC1"

/sex="male"

/issue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email :

fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 249 a 206 c 188 g 269 t 11 others

ORIGIN

Query Match 57.8%; Score 888.4; DB 9; Length 923;

Best Local Similarity 97.5%; Pred. No. 1.7e-189;

Matches 900; Conservative 10; Mismatches 12; Indels 1; Gaps

QY 570 acaggagttgacaaactggtggtgaacagtgattgacacacatcaaaaccaa--cctctgacgaca 628

Db 923 ACAGGGAGTTGACCMACTGCAAGAGTGTACACACCATCAAAACCAACCCCTGACAACA 864

QY 629 gaagaatcatcatgtgcctgttggaatccaagagatcttctctgtatggcgctgcctcat 688

Db 863 GAAGAATCWTMATGTGCGCTTGGAAATCCAGAGMTCTCTCTGTATGGSGTGCCTCAT 804

QY 689 gccatgcctctcgcagttctatgtgtgaacagtgagctgtcctgccagctgtaccaga 748

Db 803 GCCATGCCCTCTGCCAGTTCTTTTGTGTGAACAGTGWGCTGTCTTCCAGCTGTACCAGA 744

QY 749 gatcggagacatggcgctcgtgtgcctttcaacatcgcagctacgcctgtctcagt 808

Db 743 CATCGGAGACATGGGCTCGGTGTGCTTTCAACATCGCCCACTACGCKGCTAACA 684

QY 809 acatgattgcgcacatcacggcgctgaagccaggtgaactttatatacacactttggagatg 868

Db 683 ACATGATTGGCACAATCACGCGCTGAAGCCAGGTGACTTTAAACACACTTTGGGAGATG 624

QY 869 cacatattacatgaatcacatcgcagccactgaaatcagcttcagcgagaaacccagac 928

Db 623 CACATATTACCTGAATTCACATCGAGCCACTGAAAATTTCAGCTTCAGCGAGAACCCAGAC 564

QY 929 ctttcccaagctcagattcttcgaaaagttcgaataattgacttcgaactcgaagctgaag 988

Db 563 CTTTCCCAAGCTCAGGATTCCTCGAAAAGTTGAGAAAATTGACTTCAAGCTGAAG 504

QY 989 acttcagattgaaggtacacatccqcatccaactattaaaatggaaatggctgttagg 1048

Db 503 ACTTTCAGATTGAGGGTACAACTCCGATCCCACTATTAATGGAATGGCTGTTAGG 444

QY 1049 gtgccttcaagagctCgaaggtattgtcagctcttttaggggttggcgctgagtcgcag 1108

Db 443 GTGCTTTCAAAGGAGCTCGAAGCATATTTGTCAGTCTTTAGGGTTGGGTGATGCCGAG 384

QY 1109 gtaaaagtctctttctctataaagaagaagaaactaggtcaaaatctgtccgtgacct 1168

Db 383 GTAAAAGTTCTTTTCTCTAAAAGAAAAGAACTAGGTGCAAAAATCTGTCGGTACCT 324

QY 1169 atcagttatattatttttaaggatgtgcactggcaaatgtaactgtgccagttctttcc 1228

Db 323 TTCAGTTATTAATTTTAAAGGATGTTGCCACTGGCAATGTAACTGTGCGCAGTTCTTCC 264

QY 1229 ataataaagcctttgagtttaactcactcaggtggtatctgcacatcgtcaggttatgaca 1288

Db 263 ATAATAAAGGCTTTGAGTTAACTCACTCAGGCTTCTCTGACAAATGCTGAGGTATGAACA 204

QY 1289 aagtggagagaatgaaatgtatgtctcttagcaaaaaaatgtatgtgcattcaatccc 1348

Db 203 AAGTGGAGAGAATGAAATGTATGTCTCTTTAGCAAAAACAATGTTTGTGATTTCAATCCC 144

QY 1349 acgtactataaagaaggtgtgtaattcacagctatttttgaatttttagaata 1408

Db 143 ACGTACTATAAAGAGGTTGTGTGAATTCACAGCTATTTTGGAAATATTTTAGAATA 84

QY 1409 tttaagaatttcacagctattcctcctcaaatctcagggagctgagtgagtaacaccatcgc 1468

Db 83 TTTTAAAGAAATTCACAGCTATTCCTCAATCTGAGGAGCTGAGTNACACCATCGATC 24

QY 1469 atgatgtagagtggtgttatgaa 1491

Db 23 ATGATGTAGAGTGTGCTTATGAA 1


```

Qy 498 gggcccgatttgcttcacgttgaggcatttggggcagaatacacagatagatc 557
Db 1012 GGGCCCGATTATGCTTCCAGTGGAGGCAATTTGGGGCAGAAATACAGATATGGAATC 953
Qy 558 agattattcagacaggagttacacaaactgcaagagtgattgacacaaacaa 617
Db 952 AGATTATTGAGGACAGGAGTGGACCACTGCAAGAGTGATTGACACCAATCAAAACCAA 893
Qy 618 cctgcagcagaagaatcatcatgtgcgttggaatccaagagatcttctctgtgac 677
Db 892 CCCTGAGCAGACGAATCATCATGTGCGCTTGGAAATCCAGAGATCTCTCTGTGATGC 833
Qy 678 gctgcctccatgcctgcctctgcagttctatgtgtgaaacagtgagctgtctgca 737
Db 832 GCTGCTCCATGCCATGCCCTCTGCCAGTCTATGTGTGTGTAACAGTGAGCTGCTGCCA 773
Qy 738 gctgtaccagagatcgagacatggtcctcggtgtgctttcaacatcgccagctacg 797
Db 772 GCTGTACCAAGATCGGAGACATGGGCCCTCGGTGTGCTTTCAACATCGNCASPAGCC 713
Qy 798 cctgtccacgtacatgattggcacatcacgggctgcctgaagccaggtgactttatacacac 857
Db 712 CCTGTCTCACTACATGATTGGCACATCACGGGCTGAAGCCAGGTGACTTTATACACAC 653
Qy 858 ttgggagatgcacataattacatgaatcacatcgacatgcaaaattcagcttcagcg 917
Db 652 TTTGGGAGATGCACATATTTTACCTGAATCACATCGAGCCACTGAAATTCAGCTTCAGCG 593
Qy 918 agaacccagaccttcccaagctcaggtattctgaaagttagaaaattgaaattatgact 977
Db 592 AGAACCCAGACCTTCCCAAGCTCAGGATCTTCGAAAGTTGAGAAATTCATGACTT 533
Qy 978 caagctgaagacttcaagttaggttacaatccgcatccacttaaaatggaat 1037
Db 532 CAAGCTGAGACTTTCAGATTGAAGGTACATCCGATCCACTATTAAATGGAAT 473
Qy 1038 ggcgttgtaggtgcttccaagagctcgaaggtattgtcagctctttaggggttgagc 1097
Db 472 GGCCTGTTAGGTGCTTTCAAAAGGAGCTTGAAGATATTGTCACTCTTTAGGGGTGGGC 413
Qy 1098 tgatcccgaggtaaaagtcttttctcctaaagagaaggaactagtgcaaaatct 1157
Db 412 TGGATCCCGAGGTAAAGTCTTTTCTCTAAAGAAAGGAAGAACTAGTCAAAATAA 353
Qy 1158 gtcgtgacctacagttatttaagtagtggtgcactggcaaatgtaactgtgc 1217
Db 352 RTCCGTGACCTATCAGTTAATTTTAAAGGATGTGGACTGGCAATGTACTGTGC 293
Qy 1218 cagttcttccataaaaaggcttgagtttaactcactgaggtatctgacaatgtga 1277
Db 292 CAGTCTTTTCCATAATAAAGGCTTTGAGTTNACTCACTGAGGGTATCTGACAACTGA 233
Qy 1278 ggttatgaacaagtgaagagaatgaatgtatgtgctcttagcaaaaacatgtatgac 1337
Db 232 GGTATGAACAAGTGAGGAGATGAATGTATGTGCTTATGAGCAANACATGTATGTGC 173
Qy 1338 atttcaatccacgtcactataaagaaggttggtgaatttcacagctatttttgaata 1397
Db 172 ATTTCATCCACGTACTATTAAAGAAGTGTGGTGAATTTTACAGCTATTTTGGAAAT 113
Qy 1398 tttttagaatattttaagaatttcacaagctattccctcaaatctgagggagctgagtaa 1457
Db 112 TTTTGTAGATATTTTAAAGATTTTCAAGCTATTCCCTCAATCTGAGGGMGCTGAGTAA 53
Qy 1458 caccatgcataatgtagtagtggtgttatgaactttaAagttgttttata 1509
Db 52 CACCATCATGATGATGAGTGTGTTNTGAACCTTTATCGTGTGTTTATA 1

```

prime, mRNA sequence.
 AL547612.1 GI:12881832
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1020)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..1020
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DI008YH21"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cD
 was primed with a NotI-oligo(dT) primer. Five prime en
 riched, double-stranded cDNA was digested with Not I
 cloned into the Not I and Eco RV sites of the pCMVSPORT
 vector. Library was normalized. Library was constructed
 Life Technologies. Contact : Feng Liang Life Technolog
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 237 a 291 c 276 g 215 t 1 others
 ORIGIN
 Query Match 66.3%; Score 1018; DB 9; Length 1020;
 Best Local Similarity 99.8%; Pred. No. 1.2e-218;
 Matches 1018; Conservative 1; Mismatches 1; Indels 0; Gaps

RESULT 4

AL547612

LOCUS

DEFINITION

1020 bp mRNA linear EST 16-FEB-2001
 AL547612
 AL547612 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI008YH21 5

immunogenic; gene therapy; vaccine; colonic cancer; ss.
Homo sapiens.
W0200149716-A2.
12-JUL-2001.
29-DEC-2000; 2000WO-US35596.
30-DEC-1999; 99US-0476296.
10-JAN-2000; 2000US-0480321.
15-FEB-2000; 2000US-0504629.
06-MAR-2000; 2000US-0519444.
19-MAY-2000; 2000US-0575251.
29-JUN-2000; 2000US-0609448.
28-AUG-2000; 2000US-0649811.
(CORI-) CORIXA CORP.
Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
King GE, Wang T, Jiang Y;
WPI; 2001-441847/47.
Colon tumor associated proteins and nucleic acids useful for the
prevention, diagnosis and treatment of colonic cancer -
Claim 2; Page 254; 472pp; English.
The present invention describes colon tumor associated proteins (I) and
the polynucleotides (II) that encode them. (I) have cytostatic activity.
(I) and (II) can be used in gene therapy and vaccine production. (I) and
(II) may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate colon tumor associated protein (TCAP)
expression, such as colonic cancer. For example, (I) and (II) may be
used to treat disorders associated with decreased expression by
rectifying mutations or deletions in a patient's genome that affect the
activity of TCAPs by expressing inactive proteins or to supplement the
patients own production of them. Additionally, (II) may be used to
produce the TCAP proteins, by inserting the nucleic acids into a host
cell culturing the cell to express the protein. (II) and its
complementary sequences may also be used as DNA probes in diagnostic
polymerase chain reaction (PCR) and hybridisation assays to detect and
quantitate the presence of similar nucleic acids in samples, and
therefore which patients may be in need of restorative therapy. (I) may
also be used as antigens in the production of antibodies against TCAPs
and in assays to identify modulators of TCAP expression and activity.
Anti-(I) antibodies and antagonists may also be used to down regulate
TCAP expression and activity. The anti-(I) antibodies may also be used
as diagnostic agents for detecting the presence of TCAPs in samples
(e.g. by enzyme linked immunosorbant assay (ELISA)). AAI28460 to AAI29512
and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
given in the exemplification of the present invention.
Sequence 345 BP; 104 A; 79 C; 56 G; 107 T; 0 other;

Query Match 22.1%; Score 339.6; DB 22; Length 346;
Best Local Similarity 98.8%; Pred. No. 8.e-79;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 acaatccatccactataaaatggaatggctgtgttaggggtgctttcaaaaggagctc 1066
|||||
DB 346 ACAATCCGATCAACTATAAATGGAATGGCTGTGTAGGGTCTTTCAAAGAGCTT 287
|||||
QY 1067 gaaggatattgctgttttaggggttgatggatggcgaggtataaagtctcttttgc 1126
|||||
DB 286 GAAGGATATTGTCAGTCTTTAGGGGTGGGCTGGATGCCGAGGTAAAAGTCTCTTTGCT 227
|||||
QY 1127 ctaaaagaagaaggactagtcataaaatctgtccgtgacccatcagttattatttta 1186
|||||
DB 226 CTAAGAAGAAAGGACTAGTCAAAATCTGTCCGTGACCCATCAGTTATTAAATTTTA 167

QY 1187 aggatgttgcactggcgaatgtaactgtgccaagttctttccataataaaaggctttgag 1246
|||||
DB 166 AGGATGTGCCACTGGCAATGTAACTGTGCCAGTCTTTCCATAATAAAGGCTTTGAG 107
|||||
QY 1247 ttaactcactgaggggtatctgacaatgctgaggttatgaacaaagtggaggaatgaaat 1306
|||||
DB 106 TTAACCTCACTGAGGTTATCTGACAAATGCTGAGGTTATGACAAAGTGAGGAGATGAAT 47
|||||
QY 1307 gtatgtgctcttagcaaaaaacatgtatgtgcatttcaatccacgt 1352
|||||
DB 46 GTATGTGCTCTTTAGCAAAAACATGTATGTGCATTTCAATCCCAAGT 1

Search completed: July 26, 2002, 17:57:13
Job time: 12327 sec

QY	295	cgctacagcctgagagatgaattccctctgctgcacaccaaaccgctgtcttcttcggaaggggt	354
Db	20843	AGGTATAGTCTGCGGGACACTTTCCCTTATACACCAAAACGGGTGTTTGTGCGAGGC	20784
QY	355	gttttgaggaggtgctgtgtgttatcaaggagatccacaatagtcaaaagagctgtcttcc	414
Db	20783	GTAGTGCAAAGAGCTGCTTGGTWTCTGAAGGGAGTACTGCATCCAGGGAGCTTTCAGA	20724
QY	415	aaggagatgaaatctggagtgccatggatccocgagacttttbgacacgcctgggattc	474
Db	20723	ACAGGAGTCAAGATATATGGGAAAAATGGCTCCAGGGAGTTCCTGGCGGGCGGCGCTG	20664
QY	475	tcaccagagagaaggagacttggccagtttatggcttccactgagagcattttgg	534
Db	20663	GGSCACAGAAGGGGGGATTTGGGACCTGTTACGGTTCCAGTGGAGGCACTTTGGG	20604
QY	535	gcagaatacagagatatggatacagattatttcaggacagggagttgaccacactgcaaga	594
Db	20603	GGCGCGTACTGTCAGCGCGATGCTGACTATACAGGGCAGGGGTTTGACCAATTCGTAC	20544
QY	595	gtgattgcacatcaaaacacccctgacgacagaagaatcaatcattgctgcgtctggaa	654
Db	20543	ATGTGGATTTATFAAAAAATATCCGCACATAGAAGAAATCATTTATGTGTGGTGGAAC	20484
QY	655	ccaagagatcttctctgatggcgctgcctccatgcocatgcctctgccagttctctatgtg	714
Db	20483	CCGGCGACTTGTGTTGATGCGCTTCCGGCCCTGTCACTTGTATTGTCAATTTATGTA	20424
QY	715	gtgaacagtgagctgtcctgcagctgtaccagagatcgggagacatgggacctcgtgtgtg	774
Db	20423	GCTGACGCTGAGCTTTCCCTGTCAGCTGTATCAGAGTTCGGGAGACATGGGTTTGGAGTT	20364
QY	775	cttttcaacatcgccagctacgccctgctcacgtacatgattgcgcacatcaaggccctg	834
Db	20363	CTTTTACATTCGCCAGCTATTCCCTCTTACTATATAGCTTGTCTGATGTTACTGTCCT	20304
QY	835	aagccagtgactttatcacactttgggagatgcacatatttacctgaatcacatcagag	894
Db	20303	AGACCCGGGAGTTTATTCACACGTTGGAGATGCCACATCTACAAAACGCATATAGAG	20244
QY	895	ccactgaaattcaagcttcaggagaaccacagactttccaaagctcaagattcttoga	954
Db	20243	CCACTACGGCTGACGTACGCGCACTCCACGTCCTTTCCGCGCCTGGAGACTCCGG	20184
QY	955	aaagtggaaaaattgatgacttccaaagctgaagactttcagattgaaggttaacaatccg	1014
Db	20183	TCTGTTTCTTCATCGAAGATTTACACCTGATGATTTAGACTGGTGGACTACTGCCG	20124
QY	1015	catccaactattaaatgaaatgctgtttag	1047
Db	20123	CACTACCATCTGTATGAATGCAGATAG	20091
RESULT	11		
AAV19941/c			
ID	AAV19941	standard; DNA; 137507 bp.	
XX			
AC	AAV19941;		
XX			
DT	03-AUG-1998	(first entry)	
XX			
DE	KSHV	long unique coding region and terminal repeat.	
KW	KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;		
KW	interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;		
KW	complement-binding protein; glycoprotein; capsid protein IV; infection;		
KW	immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;		
KW	lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;		
XX	HIV immune status; anti-inflammatory agent; therapy; ds.		
XX	Kaposi's sarcoma-associated herpes virus.		


```

|||||
Db 181 aaatctgtccgacactatcagttatttaatttaagtgatgtgccactggcaaatgta 240
QY 1211 actgtgcagttcttccataataaaaggctttgagtttaactcaactgagggtatctgaca 1270
Db 241 actgtgcagttcttccataataaaaggctttgagtttaactcaactgagggtatctgaca 300
QY 1271 atgctgaggttatgaacaaagtggagagaatgaaatgtatgtctcttagcaaaacatg 1330
Db 301 atgctgaggttatgaacaaagtggagagaatgaaatgtatgtctcttagcaaaacatg 360
QY 1331 tatgtgatttccactcacttactataaagaaggttggtgaatttcacaaagctatttt 1390
Db 361 tatgtgatttccactcacttactataaagaaggttggtgaatttcacaaagctatttt 420
QY 1391 tggaaatattttagaatattttaaagaatttcaaaagctattccctcaaatgtgaggagc 1450
Db 421 tggaaatattttagaatattttaaagaatttcaaaagctattccctcaaatgtgaggagc 480
QY 1451 tgaataacacatcgcacgatgatgtaggtgtgtgtatgaactttaAagttgttttatat 1510
Db 481 tgaataacacacatcgcacgatgatgtaggtgtgtgtatgaactttaAagttgttttatat 540
QY 1511 tttgtctataataaagaaggtttctgc 1536
Db 541 tttgtctataataaagaaggtttctgc 566

RESULT 6
AAH74201
ID AAH74201 standard; DNA; 124884 BP.
XX
AC AAH74201;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of the Varicella virus Dumas strain.
XX
KW Dumas strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.
XX
OS Varicella virus.
XX
PN W0200156600-A1.
XX
PD 09-AUG-2001.
XX
PF 31-JAN-2001; 2001WO-JP00678.
XX
PR 31-JAN-2000; 2000JP-0062734.
XX
PA (OSAU ) UNIV OSAKA.
XX
PI Gomi Y, Sunamachi H, Takahashi M, Yamanishi K;
XX
DR WPI; 2001-48845/53.
XX
PT Method for checking absence of mutation at specific positions of
PT varicella virus genome for quality control of attenuated live varicella
PT vaccine -
XX
PS Claim 1; Page 61-105; 158pp; Japanese.
XX
CC The present sequence represents the nucleotide sequence of Varicella
CC virus, Dumas strain. The specification describes a method for the
CC quality control of an attenuated pox vaccine, characterised in that
CC the DNA sequence of specific parts of the viral genome in a sample is
CC determined and proved to be conserved rather than mutated in comparison
CC with the same parts of a reference viral genome. The method is useful
CC for quality control of vaccines for use in the prevention of poxvirus
CC diseases such as chickenpox.
XX
SQ Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 other;

```

```

Query Match      26.6%; Score 408; DB 22; Length 124884;
Best Local Similarity 66.5%; Pred. No. 2.1e-95;
Matches 585; Conservative 0; Mismatches 295; Indels 0; Gaps

QY 189 cggggaagctcagtaactgaggggagatccaaacatctcctcgtcggtcaggaagga 248
Db 18488 cggggaagctcagtaactgaggggagatccaaacatctcctcgtcggtcaggaagga 1854
QY 249 cgaccgcacgggcacccgcacccctgtcggtatttcggcatcgagggcgctacacagccgag 308
Db 18548 cgatcgacacgggaatcggaacgttatctttatttggaaatcgaaagctcgatatacaattgog 1860
QY 309 agatgaattccctctcgtcgacacccaaacgtgtgttctggaaaggggtgttttggagggagtt 368
Db 18608 aaatgaattccctctttaaactacaagcgtgttttttggagggcgctcggtggagaggtt 1866
QY 369 gctgtgtttatacaaggatccaaacaaatgctaaagagctgtcttccaagggagtgaaat 428
Db 18668 gttatgttttatacgggttcacacgattccaaagaactcgcgctaaagataatacaat 1872
QY 429 ctgggatgccaatggatcccgagactttttggacagcctgggattcttccacacagagaaga 488
Db 18728 atgggatatacaggtacgagcaaatcttcaataggaatggtcttccataaaagacacac 1878
QY 489 aggggacttggcccgagtttatgcttccagtgagggcattttggggcgagataacagaga 548
Db 18788 gggggacttggcccgagtttatgcttccagtgagggcattttggggcgagataacagaga 1884
QY 549 tatggaatcagattattcaaggagggaggttgacaaactgcaaaagtgattgtgacacacat 608
Db 18848 ctgtcaatcaactatttacagaaaggaatcgatcagctgcaaaactgttagatacaat 1890
QY 609 caaaacaaacccgtgacagacagagaatcatcatgtgcttggaaatccaaagagatcttcc 668
Db 18908 taaacaaaccccgaaagcgacgaatgattatactgttggaaatccaaagagatcttcc 1896
QY 669 tctgatggcgtcctccatgcctgcctcctgcctcctcctcctcctcctcctcctcctcctcct 728
Db 18968 cttaatggtactacctccatgacacgcttatgtcagttttacgttgcaaacggtgaatt 1902
QY 729 gtctgcagctgtacacagatcgaggagacatgggagacatgggctcgtgtgtccttccaaatcgc 788
Db 19028 atcctgcgaagtataccagagatcggggagatgggcttggggttgggtacacattgc 1908
QY 789 cagtaagcctgtcactacatgatctgacacatcacggcctgaagccaggtgactt 848
Db 19088 tggatagcactctcttaccctacatagtgagcagctgttacaggagacttaaaacccgagattt 1914
QY 849 tatcacacactttggagatgcacatatttacctgaatcacatcgacccactgaaatcca 908
Db 19148 aattcacaatgggggagtgacatatttacctgaatcacatagatgctttaaagtgca 1920
QY 909 gcttcagcgagacccagacaccttccaaagcactcaggattcttcgaaaggttgagaaat 968
Db 19208 gctagctgactcccaaaacctttccttgccttaaatatttcgaaatgttaacagatatt 1926
QY 969 tgaatcactcaaaagctgaagactttcagattgaaaggtgacaaatccgcatcccaactataa 1028
Db 19268 aaacgactttaaatggagcagatttccagcttgatggatataataatccacccccccctaaa 1932
QY 1029 aatgaaatggctgttttgggtgcttttcaaaaggagctCga 1068
Db 19328 aatgaaatggctgttttgggtgcttttcaaaaggagctCga 19367

RESULT 7
AAH74202
ID AAH74202 standard; DNA; 125157 BP.
XX
AC AAH74202;
XX
DT 15-OCT-2001 (first entry)

```


TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1292 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-283-471A-37

Query Match	3.2%; Score 48.6; DB 4;	Length 1292;
Best Local Similarity	54.1%; Pred. No. 0.002;	
Matches	99; Mismatches 84; Indels 0;	Gaps
Conservative	0;	
QY	10	ggagacaacttggcctgcgtcccgccgagcgcaacttgccctgcctccgcgcgcg 69
Ddb	443	GGGCCCCCACCGCCGGCGTGGGCCGGGGGGCGGGGCTACCCCTCCACCCCCCCTCG 502
QY	70	ccaattgcctgcctccggtccccgccgcgcgcgcacatgectgtggccgggtctcgagctg 129
Ddb	503	CGCCCCCTTCGCAGCTTCGCGCGCGCCTTCGCCCTCCGCGCTCACC GCGGAGCACTG 562
QY	130	cgcgcgcgcgccttgcctcccgcgcgcgcacagagcgaacgcgcgaccgcgtccgcac 189
Ddb	563	GCCGCGCTGCGCTGCGACGCGGGGCGGGAGGGGGCGCCGGAGCCGCCCGCGACCCCC 622
QY	190	ggg 192
Ddb	623	GG 625

RESULT 15
US-08-458-568A-3/c
: Sequence 3, Application US/08458568A
: Patent No. 5821339
: GENERAL INFORMATION:
: APPLICANT: Schafer, Priscilla A.
: APPLICANT: Yeh, Lily
: TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvir
: TITLE OF INVENTION: Infections
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339ris
: STREET: One Liberty Place, 46th floor
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Wordperfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/458,568A
: FILING DATE: 02-JUNE-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/065,146
: FILING DATE: 05-MAY-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Leary Ph.D., Kathryn R.
: REGISTRATION NUMBER: 36,317
: REFERENCE/DOCKET NUMBER: DFCI-0029
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-3100
: TELEFAX: (215) 568-3439
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 702 base pairs

Db 1668 CCAAAGAAATGGGACTTGAAGTGGTGAATTTGTCACACATTTGGAGATGCACATATT 1609
QY 877 tacttgatcaatcagccactgaataatcagcttcacgagagaccacacacacattccca 936
Db 1608 TATCAAAATCAATTTGATGCGATCGMAACACACATTTAGCAGCTGAAAGCTTCAATCTCC 1549
QY 937 aagctcagatcttcgaaagttagaataatgatgacttcaagctgaagacttt-ca 995
Db 1548 AACATTA--AAATTAACAGTGCACAAGTCTATTTTCGACATAAATTAAGAGATTGGA 1491
QY 996 gattgaaggttacaatccgcatcccaactattaaaatggaaatggctgttag 1047
Db 1490 AATTGTTGATGATGATCATCCACATCCAGCMATAAAAGCTCCMATAGCAKTGTAG 1439

RESULT 10

US-09-266-417-7/c
; Sequence 7, Application US/09266417
; Patent No. 6228588
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/266,417
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3479 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-09-266-417-7

Query Match 7.9%; Score 121.2; DB 4; Length 3479;
Best Local Similarity 56.8%; Pred. No. 1.7e-22;
Matches 268; Conservative 6; Mismatches 194; Indels 4; Gaps
QY 578 ttaccacactgcaagagtgattgacaccatcaaaacacacccctgaagacagaagaatca 637
Db 1908 TTGATCAACTTAAACAGATAATTGAACAAATTAAGCATTAATCCAGATTCAAGCGGACACA 1849
QY 638 tcctgctgccttggatccaagatcttctctgctgctgctgctgctgctgctgctgctgctg 697
Db 1848 TCGTACTCGATGGATCCAAACAGAAATGATACATGAGCCTTCCGCTTGTCAATCCA 1789
QY 698 tctgacagttctatgtgtgaacagtgagctgtcctgacagctgtaccagagatcgaggag 757
Db 1788 IGITCCAGTTTATGTCACAGATGTTAAGTTAAGTTGCGAGTTATATACCAGTAGCGCAG 1729
QY 758 acatgggctcggtgcttcaacatccagctgaagctgacccctgctgctgctgctgctgctg 817
Db 1728 ATATCTTTTAGGTGTGCCATTAATATCGARGCTACGCTTTATGACACCTTATGTG 1669
QY 818 cgcacatcag-ggctggaagccaggtgactttatacacacttttggagatgcacatt 876
Db 1668 CCAAGAAATGTGGACTTGAAGTGGTGAATTTGTGCATACATTTGGAGATGCACATATT 1609
QY 877 tactgaatcacatcagcagccactgaaatcagcttcagcagagacccacacacacacacac 936
Db 1608 TATCAAAATCAATTTGATGCGATCGMAACACACATTTAGCAGCTGAAAGCTTCAATCTCC 1549
QY 937 aagctcaggtattctcgaaagttagaataatgatgacttcaagctgaagacttt-ca 995
Db 1548 AACATTA--AAATTAACAGTGCACAAGTCTATTTTCGACATAAATTAAGAGATTGGA 1491
QY 996 gattgaaggttacaatccgcatcccaactattaaaatggaaatggctgttag 1047
Db 1490 AATTGTTGATGATGATCATCCACATCCAGCMATAAAAGCTCCMATAGCAKTGTAG 1439

RESULT 11

US-08-483-533-7
; Sequence 7, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742


```
QY 655 ccaagatctctctgagcgctgcctccatgcacatgctgcccattctatgtg 714
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20483 CCGCGGACCTGTCGTGTGATGCGCTTCGCGCTGTCACTGTGTATGTCAATTTATGTA 20424
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 715 gtgaacagtgagctgctctgcagctgtaccagagatcgggagacatgggctcggtgtg 774
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20423 GCTGACGGTGAGCTTTCCTGTCAGCTGTATCAGAGTCGCGGAGCATGGGTTGGGAGTT 20364
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 775 ccttcaacatgcagctagcctgctcaatgctgacatgctgacatcacgggctg 834
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20363 CCTTTTAACTGCGGAGCTTATTCCTCTTAACTTATATGCTTGCATGCTACTGCTT 20304
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 835 aagcagtgactttatatacacactttggagatgcacatatatttaacctgaatcacatcgag 894
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20303 AGACCCGGGAGTTTATTCACACGTTGGGAGATGCCACATCTACAAAACGATATAGAG 20244
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 895 ccaatgaaattcagcttcagcgagaaacccagacaccttcccaaaactcagattcttcca 954
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20243 CCACCTAGCGGTGAGCTGACGCGCACTCCAGTCCTCTTCCCGGCTGGAGATACTCCGG 20184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 955 aaacttgagaaattgacttcaactcaagctgaagactttcagattgaagggtacaatccg 1014
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20183 TCTGTTTCTCCATGGAAGAGTTTACACCTGATGATTTTACCTGGTGGACTACTGCCCG 20124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1015 catcaactattaaatggaatggctgtgttag 1047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20123 CATCCTACCATTCGTATGGAATGCGAGTATAG 20091
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7

```
US-08-714-918-7/c
; Sequence 7, Application US/08714918
; Patent No. 6031123
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
```

```
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
```

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
```

```
; MEDIUM TYPE: storage
```

```
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
; SOFTWARE: Word Perfect 5.1
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/714,918
```

```
; FILING DATE: September 13, 1996
```

```
; CLASSIFICATION: 424
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 60/009,102
```

```
; FILING DATE: December 22, 1995
```

```
; APPLICATION NUMBER: 60/003,798
```

```
; FILING DATE: September 15, 1995
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Warburg, Richard J.
```

```
; REGISTRATION NUMBER: 32,327
```

```
; REFERENCE/DOCKET NUMBER: 222/005
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (213) 489-1600
```

```
; TELEFAX: (213) 955-0440
```

```
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3479 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-714-918-7
```

```
Query Match 7.9%; Score 121.2; DB 3; Length 3479;
Best Local Similarity 56.8%; Pred. No. 1.7e-22;
Matches 268; Conservative 6; Mismatches 194; Indels 4; Gaps
```

```
QY 578 ttgaccactgcaaaagtgattgacaccatcaaaacccactgacacagagaatca 637
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1908 TTGATCACTTAAACAGTAATTGAACAAATTAAGCATATCCAGATTCAAGGCGACACA 1849
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 638 tcatgtcgcttggaatccaagagatcttctctgatggcgctgctccatgccatgcc 697
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1848 TCGTATCTGCATGGAATCCACAGAAATGTATACAAATGGCACTTCGGCTTGTCTATCCA 1789
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 698 tctccagttctatgtggtgaacagtgagctgtcctgccagctgtaccagagatcgggag 757
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1788 TGTTCACAGTTTATGTCCAGATGTAAGTTAAGTTGCCAGTTATACCAACGTAGCGCAG 1729
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 758 acatgggctcggtgtgcttccaacatgcacgctacgctcgcctgcctcagctacatgattg 817
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1728 ATATCTTTTGTAGTGTGCCATTTAATATGCGCARGCTACGCTTTATTTGACACACCTTATG 1669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 818 cgcacatcacg-ggctgaagccagggtgactttatatacacactttggagatgcacatatt 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1668 CCAAAGAAATGTGGACTTGAAGTGGTGAAATTTGTGCATACATTTGGAGATGCACATATT 1609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 tacctgaatcacatgcacgctgaaatgcagctcagctcagcgagaaacccagacccttccca 936
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1608 TATTCAATCATATATTCGATKCGMAACACCATTTAGCAGCTGAGAAAGCTTCAATCTCC 1549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 aagctcaggattcttcgaaaagttagaaaattgatacacttgcacttcaaaagctgaagacttt-ca 995
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1548 AACATTA--AAATTAACAGTGCACAGTCTATTTTCGACATAAAATTTATGAAGATTGTGA 1491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 gattgaaggggtacaatccgcatcccaactattaaaatggaaatggctgttttag 1047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 AATTGTTGACTATGAATCATCCACGCMATAAAAGCTCCMATAGCAKTGTAG 1439
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

```
US-09-265-315-7/c
; Sequence 7, Application US/09265315
```

```
; Patent No. 6187541
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Benton, Bret
```

```
; APPLICANT: Lee, Ving J.
```

```
; APPLICANT: Malouin, Francois
```

```
; APPLICANT: Martin, Patrick K.
```

```
; APPLICANT: Schmid, Molly B.
```

```
; APPLICANT: Sun, Dongxu
```

```
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
```

```
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
```

```
; TITLE OF INVENTION: TARGET GENES
```

```
; NUMBER OF SEQUENCES: 111
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Lyon & Lyon
```

```
; STREET: 633 West Fifth Street
```

```
; CITY: Los Angeles
```

```
; STATE: California
```

```
; COUNTRY: U.S.A.
```

```
; ZIP: 90071-2066
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
```

```
; MEDIUM TYPE: storage
```


;
;
US-09-089-195-1

[illegible]

QY 961 gagaaataatgatgcctccaaagctgaagacatttcagatigaagggtaacaatccgcatacca 1020

Db 961 GAGAAATAATGATGACCTTCAAGAGCTGAGGACTTTCAGATTGAGAGGTACAATCCGCATCCA 1020

QY 1021 actataaaatgaaatggctgtttaggtgctttccaaaggagctcGgaagatattgtcca 1080

Db 1021 ACTATATAAATGAAATGGCTGTTTAGGTGCTTTCAAAGAGCTTGAAGGATATTGTCCA 1080

QY 1081 gtcttttaggggttgggtcgatgcccagagtaaaagtctctttttgtctataaaagaagaag 1140

Db 1081 GTCTTTTAGGGGTTGGGCTGATGCCAGGTAAAAGTTCCTTTTGTCTTAAAGAAAAGG 1140

QY 1141 aactcagctcaaaaactctgcgtgacctatcagttataatttttcaaaggtattgtccact 1200

Db 1141 AACTAGGCTCAAAAATCTGCGTGACTTATCAGTTATTATTTTAAGGATGTGCCACT 1200

QY 1201 ggcataatgaactgtgccagctctttccataataaaaggctttagttaaactcactgagg 1260

Db 1201 GGCATAATGTAACGTGCCAGCTTTTCCATAATAAAAGGCTTGATGTTAACTCAGTCAGG 1260

QY 1261 gtatctgacaatgctgaagttatgaacaagtggaggagaatgaaatgtatgtccttag 1320

Db 1261 GTATCTGACAATGCTGAGGTTATGACAACAAGTGAGGAGATGAAATGATGTGCTCTTAG 1320

QY 1321 caaaaaactgtatgtgcattccaatcccaacttacttataaagaaggttgggtgaattcac 1380

Db 1321 CAAAAACAATGATGTGCAATTCATCCCACTACTTATAAAGAAGGTTGGTGAATTCCAC 1380

QY 1381 aagctattttgggaatatttttagaataatttaagaattccaagctattccctcaaat 1440

Db 1381 AAGCTATTTTGGAAATATTTTAGAATATTTTAAGAAATTTCAAGCTATTTCCTCAAAAT 1440

QY 1441 ctgaggagctgagtaacaaccatcgatcatgatgtagagttgggttgaacttaagct 1500

Db 1441 CTGAGGAGCTGAGTAACCACTCGATCATGATGTAGAGTGGTGTGAAGCTTATAGT 1500

QY 1501 tgtttattgttgcctataataaagaagttcttgc 1536

Db 1501 TGTTTTATGTGTGTATATAAAGAAGTGTTCTGC 1536

RESULT 2

US-09-318-448-11

; Sequence 11, Application US/09318448

; Patent No. 6210950

; GENERAL INFORMATION:

; APPLICANT: Johnson, William G.

; APPLICANT: Steunos, Edward S.

; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS

; FILE REFERENCE: 601-1-057

; CURRENT APPLICATION NUMBER: US/09/318,448

; CURRENT FILING DATE: 1999-05-25

; NUMBER OF SEQ. ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.0

; SEQ. ID NO. 11

; LENGTH: 18596

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-318-448-11

Query Match 39.9%; Score 613.2; DB 4; Length 18596;
Best Local Similarity 90.6%; Pred. No. 8.3e-153;
Matches 667; Conservative 0; Mismatches 63; Indels 6; Gaps

QY 807 gtacatgattggocacatcagcggcctgaagccaggtgactttatcacacatttgggaga 866
Db 15510 gaacttgttgatcacatcctctgtactgtttcacggacatgaggacattacaacag 1556

QY 867 tgcacattattacotgaatcacatcagccactgaaattcagcttcagcgagacccag 926

Db 15570 atgcacattatgcacaataatgaccttattgttttttaccttcagcgagacccag 1562

Db 1217 GSCCCTCGGAGCGGTGCGGAGAGGAGGCGCGCTGGGGA 1263

RESULT 8
AC020697/c

LOCUS
DEFINITION
Homo sapiens chromosome 18 clone RP11-145B19 map 18, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

AC020697
AC020697.3 GI:7382676
HTG: HTGS_PHASE1; HTGS_DRAFT.

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165742)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 18, clone RP11-145B19
Unpublished
2 (bases 1 to 165742)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bida, F.,
Boguslavsky, L., Boukhaltier, B., Brown, A., Burkett, G., Castle, A.,
Choepey, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearfield, K., Dewar, K., Domino, M., Doyle, M., Fenster, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, X.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, M., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (08-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On April 1, 2000 this sequence version replaced gi:6751720.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WBIR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L976
Center clone name: 145_B.19
----- Summary Statistics -----
Sequencing vector: MJ3; M77815; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160695 bases at least Q40
Consensus quality: 163157 bases at least Q30
Consensus quality: 164107 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 164942; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1339: contig of 1339 bp in length

QY 961 Ctgcgcccgccgcacaggagcgggagccgagccgggtccgcgcaggggagctcag 1020

```

join(1041. 12803,2033. 12900,5530. 5570,11049. 11134)
13449. 13624,14133. 14204,15613. 15750)
/EC_number="2.1.1.45"
/codon_start=1

```

/product="thymidylate synthase"


```
||||| 557 AAATGCAAAATCCCTTATTAGTTGTAGGAACACAGATCTCAACAGCAGTTTGTGACAAG 616
||||| 541 accgcagaaacagtggaactgtgtctgtctgttagaagagcggtcgaccagcgg 600
||||| 617 ACCGCAGAAAACGTTGGAACTGTGCTGTGCTTAGAAGAGCGCGTCCACAGACGG 676
||||| 601 ttcccaaggcgagctcctccengccacgcacactgontccaggttcccggtntcct 660
||||| 677 TTCCCAAGGGCGCAGCTCTTCCAGCCACCGACCTGCATCCAGTTCCCGGGTTTCCT 736
||||| 661 aaagctcagctgtggccctgggctccgttctgtccacacccctgctcctgtc 720
||||| 737 AAGACTCTACCTGTGCGCTTGGGCTCCGTCTGTGCGACCCGTGCTCCTCGTTTC 796
||||| 721 cccctggcgacgctctctctagagcgggcgccgcgcgcaccccgagcaggaagcg 780
||||| 797 CCCCTGCGCAGCCTCTCTAGAGCGGGGGCGCGCGACCCCGCGACGAGAGGGCG 856
||||| 781 gagcgcgagcggcgccgggaaagcgcggaaggggttctctgcaccgcgcaacttg 840
||||| 857 GAGCGGGGACGCGCGCGGAAAGGCGCGGAAGGGTCTCTGCCACCGCGCCACTTGG 916
||||| 841 cctgctcctgctcccgccgcgcaacttagcctgctcctcctccgcgcgcaacttcg 900
||||| 917 CTTGCTTCCGTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 976
||||| 901 cctgctcccgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
||||| 977 CTTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1036
||||| 961 Ctgcgcccgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020
||||| 1037 TTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
||||| 1021 taccctggcgagatcaacacatcctcgtggtggtggtggtggtggtggtggtggt 1080
||||| 1097 TACCTGGGCGAGATCACACATCTCCGCTGCGGCGGTGAGGAGGAGCGCGCGGCG 1156
||||| 1081 accgcaacctgcggtattcgccatgcagcgcgctgctagcctgagaggtgagcgcg 1140
||||| 1157 ACCGGCACCTCTGCGTATTCGGCATGCGCGCGCGCGCGCGCGCGCGCGCGCG 1216
||||| 1141 ggcctctcgagcgggtggtggtggtggtggtggtggtggtggtggtggtggtg 1187
||||| 1217 GCGCCCTGCGGCGCGGCGGCGGAGGAGGAGGAGGCGCGCGCGCGCGCGCGCG 1263

RESULT 4
AX30682 AX30682 18596 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 1191 from Patent WO0194629.
DEFINITION AX30682
ACCESSION AX30682
VERSION AX30682.1 GI:18103660
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (sites)
Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.K. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 1191 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
```


[illegible]

Query Match 98.0%; Score 1163.8; DB 6; Length 18596;

Db 19681 ttccacctgttggtcaggtggtctcgaactcctgacctcaggtgattccaccgcgcttgg 19740
QY 356 ccccccagaatcattacagctgacacgcgtccagcgcctcagcgcctggcgtttt 414
Db 19741 cctcccaagtgctgggtacagcgtgagtcactgaccccgccaccacagcctt 19799

RESULT 15
ABAL16598/c
ID ABAL16598 standard; DNA; 10820 BP.
XX
AC ABAL16598;
XX
XX
XX 23-JAN-2002 (first entry)
XX Human nervous system related polynucleotide SEQ ID NO 8929.
XX
XX Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0185350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.


```

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis.
XX
XX Disclosure; SEQ ID NO 36572; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 56743 BP; 14298 A; 13270 C; 14240 G; 14935 T; 0 other;

Query Match 17.2%; Score 204.2; DB 22; Length 56743;
Best Local Similarity 71.3%; Pred. No. 1.2e-26;
Matches 320; Conservative 0; Mismatches 105; Indels 24; Gaps 3

QY 1 gatcgccgcaactgcactcagcctgggtgagagagcagcagctctgtctcaaaaaaa 60
Db 1834 gattgtgcccactgcactcagcctgggtgagagagcagcagctctgtctcaaaaaaa 1893
QY 61 aaaaagacgcgcagggttcaaaacaaacccctcgaaagccctggcgctctttttt 120
Db 1894 taaataatgaaaaactcagtcgagactcagtcagctctgtctcttctcttcttct 1949
QY 121 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt 180
Db 1950 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt 2009
QY 181 cggatcttggtcactgcacactctgcctccacaggttcaagcaattctctcctcagcc 240
Db 2010 gtgatctcgctcactgcagcctctgcctcctcctcctcctcctcctcctcctcagcc 2069
QY 241 tcccaagtag-----caccacgcgcagcagcagcagcagcagcagcagcagcagc 281
Db 2070 tcccgagtagctggaattacaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2129
QY 282 gtgagacgggggtttcaccactgttgcaggtgtgtctngaacctcctcagcctcaggtga 341
Db 2130 gcagagac-ggggtttctctcttctgttggccagggtgtgtctcgaactcgtcctcaggtga 2188

```

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23014.
XX Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS W0200157182-A2.
XX
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225287.
PR 14-AUG-2000; 2000US-0225288.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0222081.
PR 12-SEP-2000; 2000US-0231958.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54342 to AAK34950 and AAK62169 represent sequences used in the exemplification of the present invention.

Query Match	17.2%; Score 204.6; DB 22; Length 11553;
Best Local Similarity	72.0%; Pred. No. 9,8e-27;
Matches 316; Conservative	0; Mismatches 101; Indels 22; Gaps 3;
QY	1 gatcgccactgcactccagcctcgggtgagagcgagactctgtctcaaaaaaaa 60
DB	8023 GATCGGCCACCGACTCCAGCTCGGCGACACTGCAGCTCGCTCAAAAAATAA 7964
QY	61 aaaaagaccgcaggggtcacaacaa--aaacctcgaaaagcctggcggtcttttt 118
DB	7963 AATAAATAAATAAATAAATAAATAAATAAATAAAGTCAGCTGATGACAATAG 7904
QY	119 ttt 178
DB	7903 CTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 7844
QY	179 gtcggtatcttggtcactgcacactctcctccaggtttcaagcaattttctgcctcag 238
DB	7843 GTGGATCTCAGCTCACTGCAACTCTGCCTCTGTGATTCAGTGTATCTCTGCTCAG 7784
QY	239 cctcccaagtag-----ccaccagccccagctcaatttttgtanttt 279
DB	7783 TCTCCTGAATAGTGGATTACAGATGTGCACACCACACCCCGCTAAATTTTGCAATTT 7724
QY	280 tagtaagacgggggtttacacctttgtccagctggtctgttgaactcctgaactcaggt 339
DB	7723 TAGTAGAGA-TGGGGTTTTGCCATGTGGCCAGCGCTGCTTTGAACCTCCAGACCTCAAGT 7665
QY	340 gatcacccgcctcggcccccnaaagtactaggtattacagcggtgagccacggctcacg 399
DB	7664 GATTTGCTCCCTCTGGCTCCCAAAAGTCTGGGATTACAGCGCTGAGCACCACGCCGG 7605
QY	400 cgcctcggcggtttttaat 418
DB	7604 CCAACAATCAGCTTTTGATT 7586

RESULT	10	
AAK66874/C		
ID	AAK66874	standard; DNA; 37449 BP.
XX		
AC	AAK66874;	
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human	immune/haematopoietic antigen genomic sequence SEQ ID NO:21686.
XX		
KW	Human;	immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX		
XX	cystatic;	gene therapy; vaccine; metastasis; ds.
OS	Homo sapiens.	
xx		
xx		
PN	WO200157182-A2.	
XX		

CC Cancers and cancer metastases of haematopoietic-derived cells, especially
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK8216
XX represent sequences used in the exemplification of the present invention
SQ Sequence 11453 BP; 3521 A; 2392 C; 2561 G; 2979 T; 0 other;

Query Match 17.2%; Score 204.6; DB 22; Length 11453;
Best Local Similarity 72.0%; Pred. No. 9.8e-27;
Matches 316; Conservative 0; Mismatches 101; Indels 22; Gaps

QY 1 gatcgccactgacgtccagctgggtgagagcgagactctgtctcaaaaaaa 60
Db 7923 GATCGCGCCACCGCCTCCAGCTGGGCGACAGACTCCGCTCAAAAAATFAA 7864

QY 61 aaaaagaccgcccgggtcacaacaa--aaacctcgaaaagccctggcggtcttttt 118
Db 7863 AATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 7804

QY 119 ttt 178
Db 7803 CTT 7744

QY 179 gtcggatcttggtcactgcaacctctgctccaggttccaggaattcttctgcctcag 238
Db 7743 GTGCGATCTCAGCTCAGTCAACCTCTGCTCTGCTGCTGCTGCTGCTGCTGCT 7684

QY 239 cctcccaagtag-----ccaccagccagcctaattttgtanttt 279
Db 7683 TCTCTGATAGTAGTGGGATACAGATGTCACACCCACCCGCGCTAATTTTGCATTT 7624

QY 280 tagtagagacgggggtttccacatgttgcaggtgctgctgaactctgacctcaggt 339
Db 7623 TAGTAGAGA-TGGGGTTTGGCATGTTGCCAGGCTGGTGTGAACTCCAGACCTCAAGT 7565

QY 340 gatccaccgctcggcccccacaaagtagtactaggtacaggtgagccacggcgtccag 399
Db 7564 GATTGTGTCCTCCCTGGCCCTCCAAAAGTCTGGGATACAGCGCTGAGCCACGCCCGG 7505

QY 400 cgcctcgcggttttaatt 418
Db 7504 CCAACAATCAGCTTTGATT 7486

RESULT 9
AAK73271/c
ID AAK73271 standard; DNA; 11553 BP.
XX AC AAK73271;
XX AC AAK73271;
DT 06-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28083.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.

PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.

Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -
Disclosure; SEQ ID NO 28084; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
supplement the patients own production of (I). Additionally, (I)
polynucleotides may be used to produce the secreted (I), by inserting
the nucleic acids into a host cell and culturing the cell to express the

PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0241617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 DR
 XX

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.
 PS Disclosure; SEQ ID NO 20393; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.
 XX Sequence 56632 BP; 17016 A; 8910 C; 10552 G; 20154 T; 0 other;
 SQ

Db	184	GTACCTGGGGCAGATCCAACACATCTCTCCGTCGGCGGTTCAGAAAGGACGACGCGG	243
QY	1080	cacgggcacccctgtcgttattcggcgatcgaggcgctacacgcctgagagtgta	1133
Db	244	CACGGGACCCCTGTCGTATTCCGATCGCATCGAGGCGCTACAGCCTGAGAGATGA	297
RESULT	14		
AL542409			
LOCUS			
DEFINITION	AL542409 LTI_FL002_P11 Homo sapiens cdna clone CS0E010YH04 5 prime	935 bp mRNA linear	EST 16-FEB-2001
ACCESSION	AL542409		
VERSION	AL542409.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 935)		
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope		
	Genoscope - Centre National de Sequenage		
	BP 191 91006 EVRY cedex - France		
	Email: segrefgenoscope.cns.fr, Web : www.genoscope.cns.fr.		

```

FEATURES
source
1..935
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE010YH04"
/clone.lib="LTI_FL002_PL1"
/lab.host="DH10B"
/note="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(gt) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
208 a 279 c 254 g 190 t 4 others
BASE COUNT
ORIGIN

```

Query Match	23.6%; Score 280.2; DB 9;	Length 935;
Best Local Similarity	98.2%; Pred. No. 1.5e-24;	
Matches 280;	Conservative 3; Mismatches 2;	Indels 0; Gaps 0;

QY	849	cgtccgcgcgccaccttgccctcgtccgcgcggcccaacttcgcctgcctccgtc	908
Dd			
	1	CGTCCCGCGCGCACTTGCGCTCCGTCGCCCTCCGCCTCCGCGCACACTTGCGCTGCGCTCCGTC	60
QY	909	cccgcgccgcgcgcacatgccttgcgcgcgtcgagagtccgcgcgcgcgcacctgcccc	968
Dd			
	61	CCCCGGCCGCGCGCATGTGCGGTGTCGGAGCTGCGCGCGCGCGCCTTGCGCCCC	120
QY	969	cgcgcacaggagcgggacgcgcgagcgcgtccgcgcgcagcgggagctgcagtaacctggg	1028
Dd			
	121	CGCCGCACAGGACCGGACGCGCGAGCGCGTCCGCCGCACGGGGAGATGCAGTAACCTGGG	180
QY	1029	gcagatccaacacatcctccgctcgtcggcgctcgaaggagcacgcgcgcgcgcaccggcacc	1088
Dd			
	181	GCAGATCCAACAACATCCTCGCTCGCTGCGCGTGCAGGAAGGACGAOCGCGACGGGACACGGCAC	240
QY	1089	ccctgcggtattcgcgatcgagcgcgscctacagcctgagagttga	1133
Dd			
	241	CCGTGCGGTATTTCGCATGCAAGCGCGCTACAGCCTGARATGCA	285
Dd			

RESULT_15	524 bp	mRNA	linear	EST 21-MAR-2001
BG4622405	602574658F1	NTH_MGC_21	Homo sapiens	cdna clone IMAGE:4650280 5',
LOCUS	mRNA sequence.			
DEFINITION	BG482405			
ACCESSION	BG482405			
VERSION	BG482405.1			
KEYWORDS	GI:13414684			
EST.				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 524)			
TITLE	NTH-MGC http://mgc.nci.nih.gov/ .			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999).			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabbs-r@mail.nih.gov			
	Tissue procurement: ATCC			
	cDNA Library Preparation: Ling Hong/Rubin Laboratory			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLCMI431 row: g column: 17			
	High quality sequence stop: 524.			

```

Query quantity sequence: 524.
location/Qualifiers
    l..524
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4650280"
    /clone_lib="NIH_MGC_21"
    /tissue_type="choriocarcinoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ; placenta; Vector: pOHB7; site: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      90 a   177 c   162 g   95 t
ORIGIN
Query Match          23.6%; Score 279.6; DB 10; Length 524;
Best Local Similarity 96.0%; Pred. No. 2.4e-24;
Matches 309; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

QY      812 ggaagggttcctgcaacgcggccacttgctcctcgctccgcgcgcgcacctggcct 871
Db       2 gGAAGGGGTCTTGCCACC CGCGCACATWGCCCTGCGCTCGCTCCGCCGCGCCACTTGGCCT 61

QY      872 gcctccgctccgcgcgcaccttgctcctcgctccccgcgccgcgcgcattgcctg 931
Db       62 GCCTCGTCCGCGCGGCACATTGCGCTGCGCTTCGTTCCCTCCGCGCGCGCG -GCATTGCCCTG 120

QY      932 ttgcgcggtcgtggagtgcgcgcgcgcgcgcctgcccctccgcgcgcacaggagcgggacgcgc 991
Db     121 TGCGCGGCTCGGA---CTGCGCGCGCGGCTTGCCCTCCCGCGGCACAGGAGCGGACGCGC 177

QY      992 agccgcgtccgcgcgcgggagctgcagttaactctgggcgcagatccaacaactctccgct 1051
Db     178 AGCGCGTCCGCGCACGGGGAGCTCATAGTACCTGGGGCAAGATCCAACAACATPCTCCGCT 237

QY     1052 cgccgctcaggaagcacgcgcgcgcgcgcgcacctgctcggtattcgcgatgcagg 1111
Db     238 GCGCGCTCAGGAGGACGACCGGACCGGCGACCGGCAACCCCTGTCGATTCGCGATGCGAGG 297

QY     1112 ccgcgtacagcctgagaggtga ll33

```

